

QY 463 GAGAAGATTATGAGGATCGCAAGCCAGGTCCTTGGTGTCTCAAACTGGACCATGGCC 522
|||||
Db 5669 GAGAAGTCTAAGATGACGAGATTGGCCAGTCCATCGGGTGTCCAATTCACACACAGG 5728
|||||
QY 523 GACCTTGAGAGATGTCCTCAAGTTGCG-----CAAGGTCACTGCTCAGCCCAACGACATC 576
|||||
Db 5729 CTGTGGAGATGATCTCTCAACAGCCAGGGCTCAAGTACAGGCTGTCTGCAACACAGGTG 5788
|||||
QY 577 GAGATTACCCCTTCTCTGCCCAACAGGAGGTGGTGCAGTACTCTCTTCTCCAAAGAACATT 636
|||||
Db 5789 GAATGTCACTTACTTCAACAGAGAAACTGCTGGATTTCTGCAAGTCAAAAGACATT 5848
|||||
QY 637 ATGCCCGTGGCTACTCTCTCTGGGTGCGAGAACAGAGTTCC-----CACACC 687
|||||
Db 5849 GTTCTGGTGTCTATAGTGTCTCTGGATCCATCGAGAGAACCAATCGAGTGGACCCGAAC 5908
|||||
QY 688 GGTGAGCGGTTCAGGAGAACAGACTCTGAACAGATCGCGAGAGGGGGGCAACACC 747
|||||
Db 5909 TCCCGGTGCTCTGGAGAACAGTCTTGTGCTTGGCAAAAGACACAGCGAAC 5968
|||||
QY 748 CTGTCTAGGTTCTTATTGCTGGGTCTGGCGCTGGCTACGTCTGTTCTCCCAAGAGC 807
|||||
Db 5969 CCAGCCCTGATTGCCCTCGTACCAGCTGACGGTGGGGTGTGGTCTCTGGCCCAAGAGC 6028
|||||
QY 808 TCCACCCCAAGCGATTGAGTCCAACTTCAGAGCATTGAGCTC 852
|||||
Db 6029 TACATGAGCGCGCATCAGACAGAACGTGCAGGTGTTGAATTC 6073
|||||

RESULT 15

US-09-614-124B-784
; Sequence 784, Application US/09614124B
; Patent No. 630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 784
; LENGTH: 6353
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-784

Query Match 10.9%; Score 106.6; DB 4; Length 6353;
Best Local Similarity 50.5%; Pred. No. 7e-21;
Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;
QY 103 GCTGTCAACACTGCCCTCAAGACCGTTACCGTCACTTGGACTGTGCTGTACTACTG 162
Db 5324 GCCGTCAAATTGGCAATAGAACCGGGTTCACCATATTGATTCTGCACATGTTTACAAT 5383
QY 163 AACGAGGGTGGGTTGGTGGGTTATCCGTGACTTCTCTGAAGAGAACCCCTCGGTGAAG 222
Db 5384 AATGAGGAGCAGTTGGACTGGCCATCG---AAGCAGATTGCAGATGGCAGTGTGAAG 5440
QY 223 CGTGAGGACATCTTGGTCTGACCAAGGTGTGGAACCACTCCACCGTTATGAGGACGTC 282
Db 5441 AGAGAAGACATATTCTACACTTCAAGCTTTGGAGCAATTCCTCATCGACAGAGTTGGTC 5500
QY 283 CTCTGGTCCATTGACGACTCCCTGAAGCGTCTTGGACTGTGACTAGTGTGATGTTCTCTC 342
|||||

Search completed: May 23, 2004, 02:42:24
Job time : 81 secs

Db 5501 CGACCAGCCTTGGAAAAGGTCACTGAAAAATCTTCAATTGGACTATGTTGACCTCTATCTT 5560
QY 343 GTTCACTGCGCCCATTTGCTGCCGAGAAGATGCGCAGGGTGAGCCCAAGATTGGCCCTGAC 402
Db 5561 ATTCAATTTCCAGTGTCTGTAAAGCCAGGTGAGGAG-----TGATCCCAAA 5608
QY 403 GGCAAATACGTCAATTTCTCAAGGACCTGACCCGAGAACCCCGAGCCACATGCGCGCTATG 462
Db 5609 GATGAAAATGGAATAATACTATTTCACACAGTGGATCTCTGTGCCACATGGAGGCCATG 5668
QY 463 GAGAGATTATTAGGATCGCAAGCGCAGGTCCATTGGTGTCTCCAACTGGACCATTTGCC 522
Db 5669 GAGAATGTAAAGATGACGAGATTGCCAAGTCCATCGGGGTGTCCAATTTCAACACAGG 5728
QY 523 GACCTTGAAGAGATGTCCAAGTTTCGC-----CAAGGTCACTGCTCAGCCCAACAGATC 576
Db 5729 CTGCTGGAGATGATCTCTCAACAAGCCAGGGCTCAAGTACAAGCCTGTCTGCAACCCAGGTG 5788
QY 577 GAGATTCAACCCCTTCTCTGCCCAACGAGGAGCTGGTGTGCTGCTCTCCAGAGACATT 636
Db 5789 GAATGTCACTCTTACTTCAACACAGAGAAACTGCTGGATTTCTGCAAGTCAAAAGACATT 5848
QY 637 ATGCCCGTGGCTTACTCTCTCTGGGTCTCGCAGAACCCAGGTTCC-----CACACC 687
Db 5849 GTTCTGGTTGCCCTATAGTGTCTCTGGATCCCATCGAGAGAACCATGGGTGACCCGAAC 5908
QY 688 GGTGAGCGGTTCAGCGAGAAACAGACTCTGAACAGAGATCGCCGAGAGAGGGGGCAACACC 747
Db 5909 TCCCGGTGCTCTTGGAGGAGCCAGTCTCTTGTGCTTGGCAAAAGACACAAAGCGAAC 5968
QY 748 CTGTCTCAGGTTCTTATTGCTGGGTCTGCGCGCTGGCTAGCTGTTCTCCCAAGAGC 807
Db 5969 CCAGCCCTGATTGCGCTTCCGCTACCACTGACGCTGGGGTGTGGTCTCTGCGCAAGAGC 6028
QY 808 TCCAAACCCCAAGCGATTGAGTCCAACTTCAAGACATTGAGCTC 852
Db 6029 TACATGAGCAGCGCATCAGACAGAACGTGCAGGTGTTGAATTC 6073
|||||

```
QY 808 TCCAAACCCAGCGCAATTGAGTCCAACTTCAAGAGCAATTGAGCTC 852
DB 1145 TACATGAGCAGCGCATCAGACAGAACGTCGAGGTGTTGATTC 1189

RESULT 13
US-09-702-705-784
; Sequence 784, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 784
; LENGTH: 6353
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-784

Query Match 10.9%; Score 106.6; DB 4; Length 6353;
Best Local Similarity 50.5%; Pred. No. 7e-21;
Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

QY 103 GCTGTCAACCTGCCCTGAAGACCGGTTACCGTCACITGGACTGTGCTGTACTACCTG 162
DB 5324 GCGGTCAATTGGCAATAGAGCGGGTCCACCATAATTGATCTGCACATGTTTACAAT 5383
QY 163 AACGAGGTGAGTGTGGTGAAGGCTTCCGTGACITTCCTGGAAGGAGAACCCCTCGGTGAAG 222
DB 5384 ATGAGGAGCAGTGTGGTGAAGGCTTCCGTCATCG---AAGCAAGATTGCGAGTGTGAAG 5440
QY 223 CTTGTGTCATTCAGCTTCCCTGAAGGCTTCCGTGACITGGACTGTGCTGTACTGTCTCTC 342
DB 5501 CGACGAGCTTGAAGGCTCACTGNAATAATTCTCAATTGGACTATGTTGACCTCTATCTT 5560
QY 343 GTTCACTGGCCCAATGCTGCCGAGAGAAATGGCCAGGCTGAGCCCAAGATTGGCCCTGAC 402
DB 5561 ATTCAATTTCCAGTGTCTGTAAAGCCAGGTGAGGAAG-----TGATCCCAAAA 5608
QY 403 GCGCAATACGTCTTCTCAAGGACCTGACCGAGAACCCGAGCCACATGCGCGGTATG 462
DB 5609 GATGAAATGGAATAATCTATTTGACACAGTGGATCTCTGTGCAATGGAGGCGCATG 5668
QY 463 GAGAGATTTATGAGGATCGCAAGGCCAGGTCATTTGGTGTCTCCAACTGGACCAATGCC 522
DB 5669 GAGAGTGTAAAGATGACAGGATTTGGCCAAAGTCCATCGGGGTGTCCTCAACTTCAACACAGG 5728
QY 523 GACCTTGAGAGATGTCGAAGTTGCG-----CAGGTCACTGCTCAGCCCAACCATC 576
DB 5729 CTGCTGGAGATGATCCTCAACAGCCAGGCTGCAAGTCAAGAGCTGTCTGCAACAGGCTG 5788
QY 577 GAGATTTCAACCTTCTCTGCCCAACGAGGAGCTGTGTGACAGTACTGCTTCTCCAAAGACATT 636
DB 5789 GAAATGATCTCTTACTTCAACAGAGAAATCTGTGGATTTCTGCAAGTCAAGACATT 5848

RESULT 14
US-09-736-457-784
; Sequence 784, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 784
; LENGTH: 6353
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-784

Query Match 10.9%; Score 106.6; DB 4; Length 6353;
Best Local Similarity 50.5%; Pred. No. 7e-21;
Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

QY 103 GCTGTCAACCTGCCCTGAAGACCGGTTACCGTCACITGGACTGTGCTGTACTACCTG 162
DB 5324 GCGGTCAATTGGCAATAGAGCGGGTCCACCATAATTGATCTGCACATGTTTACAAT 5383
QY 163 AACGAGGTGAGTGTGGTGAAGGCTTCCGTGACITTCCTGGAAGGAGAACCCCTCGGTGAAG 222
DB 5384 ATGAGGAGCAGTGTGGTGAAGGCTTCCGTCATCG---AAGCAAGATTGCGAGTGTGAAG 5440
QY 223 CTTGTGTCATTCAGCTTCCCTGAAGGCTTCCGTGACITGGACTGTGCTGTACTGTCTCTC 342
DB 5501 CGACGAGCTTGAAGGCTCACTGNAATAATTCTCAATTGGACTATGTTGACCTCTATCTT 5560
QY 343 GTTCACTGGCCCAATGCTGCCGAGAGAAATGGCCAGGCTGAGCCCAAGATTGGCCCTGAC 402
DB 5561 ATTCAATTTCCAGTGTCTGTAAAGCCAGGTGAGGAAG-----TGATCCCAAAA 5608
QY 403 GCGCAATACGTCTTCTCAAGGACCTGACCGAGAACCCGAGCCACATGCGCGGTATG 462
DB 5609 GATGAAATGGAATAATCTATTTGACACAGTGGATCTCTGTGCAATGGAGGCGCATG 5668
QY 463 GAGAGATTTATGAGGATCGCAAGGCCAGGTCATTTGGTGTCTCCAACTGGACCAATGCC 522
DB 5669 GAGAGTGTAAAGATGACAGGATTTGGCCAAAGTCCATCGGGGTGTCCTCAACTTCAACACAGG 5728
QY 523 GACCTTGAGAGATGTCGAAGTTGCG-----CAGGTCACTGCTCAGCCCAACCATC 576
DB 5729 CTGCTGGAGATGATCCTCAACAGCCAGGCTGCAAGTCAAGAGCTGTCTGCAACAGGCTG 5788
QY 577 GAGATTTCAACCTTCTCTGCCCAACGAGGAGCTGTGTGACAGTACTGCTTCTCCAAAGACATT 636
DB 5789 GAAATGATCTCTTACTTCAACAGAGAAATCTGTGGATTTCTGCAAGTCAAGACATT 5848
```


Db 500 AATGAGGAGCAGGTTGGAGTGGCCATCCG---AAGCAAGATTGCAGATGGCAGTGTGAAG 556
Qy 223 COTGAGGACATCTTCGTCTGACCAAGGTGTGGAAACCACTCCACCGTTATGAGGAGCTC 282
Db 557 AGAAGAGACATATCTACATTCAAAGCTTTGGAGCAATCCCATCGACAGAGTTGGTC 616
Qy 283 CTCTGGTCCATTGACGACTCCCTGAAGCGTCTTGAGACTTGACTAGTTGATATGTTCTC 342
Db 617 CGACCAAGCCTTGAAAGGTCACTGAAAAATCTTCAATTGGACTATGTTGACCTCTATCTT 676
Qy 343 GTTCACTGGCCATCTGCGCAGAGAAATGCGCAGGTTGAGCCCAAGATTGGCCCTGAC 402
Db 677 ATTCAATTTTCCAGTCTGTAAAGCCAGGTGAGGAG-----TGATCCCAAAA 724
Qy 403 GGCATAATAGTCAATCTCAAGGACCTGACCGAGAACCCCGAGCCACATGCGCGCTATG 462
Db 725 GATGAAATGGAAAAATACTATTGACACAGTGGATCTCTGTGCCACATGGGAGGCCATG 784
Qy 577 GAGATTACCCCTTCTGCGCCAAAGGAGGTGTGAGTACTGTTCTTCCAAAGCAATTTGCC 636
Db 905 GAATGTCATCTTACTTCAACCAGAGAAAACTGTCTGGATTTCTGCAAGTCAAAAGACATT 964
Qy 637 ATGCCCGTGGCTACTCTCTCTGGGCTCGCAGAACCAAGTTCC-----CACCACC 687
Db 965 GTCTGGTGGCTATAGTCTCTGGATCCCATCGAGAAAGACCTGAGTGTGGTCTGCGCAAGC 1024
Qy 688 GGTGAGCGGGTCAGGAGAACAAAGACTCTGAACGAGATCGCGAGAGGGCGGCAACACC 747
Db 1025 TCCCGGTGCTCTTGAGGACCCAGTCTTTGTGCTTGGCAAAAAGCACAAGCGAAC 1084
Qy 748 CTGTCTAGGTTCTTATTGCGTGGGCTGCGCGGTGCTAGTCTGTTCTTCCCAAGAGC 807
Db 1085 CCAGCCCTGATTGCGCTACGCTACGAGTGGGTTGTGGTCTGCGCAAGAGC 1144

RESULT 10

US-09-614-124B-807
; Sequence 807, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 807
; LENGTH: 3829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-807

Query Match 10.9%; Score 106.6; DB 4; Length 3829;
Best Local Similarity 50.5%; Pred. No. 5.5e-21;
Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;
Qy 103 GCTGTCCACCACTGCCCTGAAGACCGGTTACCGTCACTTGGACTGTGCTGTACTACTCTG 162
Db 440 GCGTCAAAATTGGCAATAGAGCGGGTTCACCAATATTGATTTGCACATCTTTACAAT 499
Qy 163 AACAGCGGTAGGTTGGTGAAGGTATCCGTGACTTCTCTGAAGAGAACCCCTCGGTGAAG 222
Db 500 AATGAGGAGCAGGTTGGACTGGCCATCCG---AAGCAAGATTGCAGATGGCAGTGTGAAG 556
Qy 223 CGTGAGGACATCTCTCTGCAACGAGGTGGAACCACTCCACGTTATGAGGAGCTC 282
Db 557 AGAAGAGACATATCTTACACTTCAAAAGCTTTGGAGCAATCCCATCGACAGAGTTGGTC 616
Qy 283 CTCTGGTCCATTGACGACTCCCTGAAGCGTCTTGGACTTTGACTAGTTGATATGTTCTC 342
Db 617 CGACCAAGCCTTGAAAGGTCACTGAAAAATCTTCAATTGGACTATGTTGACCTCTATCTT 676
Qy 343 GTTCACTGGCCATCTGCGCAGAGAAATGCGCAGGTTGAGCCCAAGATTGGCCCTGAC 402
Db 677 ATTCAATTTTCCAGTCTGTAAAGCCAGGTGAGGAG-----TGATCCCAAAA 724
Qy 403 GGCATAATAGTCAATCTCAAGGACCTGACCGAGAACCCCGAGCCACATGCGCGCTATG 462
Db 725 GATGAAATGGAAAAATACTATTGACACAGTGGATCTCTGTGCCACATGGGAGGCCATG 784
Qy 463 GAGAAATTTATGAGATCGAAGGCCAGGTCAATTTGGTGTCTTCAACTGAGACCAATTGCC 522
Db 785 GAGAACTGTAAAGATCGAGGATTGGCCAAAGTCCATCGGGGTGTCCAACCTTCAACCACAG 844
Qy 523 GACTTGAAGATGTCCTCAAGTTCCG-----CAGGTCATGCTCAGCCCAACCAAGATC 576
Db 845 CTCTGGATGATCTCTCAAGCCAGGCTCAAGTCAAGCCTGTCTGCAACCAAGTG 904
Qy 577 GAGATTACCCCTTCTGCGCCAAAGGAGCTGGTSCAGTACTGTTCTTCCAAAGCAATTTGCC 636
Db 905 GAATGTCATCTTACTTCAACCAGAGAAAACTGTCTGGATTTCTGCAAGTCAAAAGACATT 964
Qy 637 ATGCCCGTGGCTACTCTCTCTGGGCTCGCAGAACCAAGTTCC-----CACCACC 687
Db 965 GTTCTGGTGGCTATAGTCTCTGGATCCCATCGAGAAAGACCTGAGTGTGGTGGACCGAAC 1024
Qy 688 GGTGAGCGGGTCAGCGAGAACAAAGACTCTGAACGAGATCGCGAGAGGGCGGCAACACC 747
Db 1025 TCCCGGTGCTCTTGAGGACCCAGTCTTTGTGCTTGGCAAAAAGCACAAGCGAAC 1084
Qy 748 CTGTCTAGGTTCTTATTGCGTGGGCTGCGCGGTGCTAGTCTGTTCTTCCCAAGAGC 807
Db 1085 CCAGCCCTGATTGCGCTACGCTACGAGTGGGTTGTGGTCTGCGCAAGAGC 1144
Qy 808 TCCAAACCCCAAGCGCAATTGAGTCCCAACTTCAAGAGCAATTGAGCTC 852
Db 1145 TACATGAGCAGCGCATCAGACAGACGTCGAGGTGTTTGAATTC 1189

RESULT 11

US-09-671-325-807
; Sequence 807, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

Qy	523	GACCTTGAGAGATGTC	CCAAATTGCG-----CAAGGTCAATGCTTCAGCCCAACGAGATC	576
Db	796	CTGCTGAGATGATCCTT	CAACAAGCCAGGGCTCAAGTACAAGCTGTCTGCAACGAGGTG	855
Qy	577	GAGATTCAACCCCTTCCT	CGCCCAACGAGGAGCTGTGTGCAGTACTGTTCTCTCAAGAACATT	636
Db	856	GAATGTCATCCTTACTT	CAACCGAGAGAAACTGCTGGATTCTGCAAGTCAAAAGACATT	915
Qy	637	ATGCCCTGCGCTACTCT	CTCTCTGGGCTCGAGAACCGAGTTCC-----CACCACC	687
Db	916	GTTCTGTTGCCATATAG	TGTCTGGGATCCCATCGAAGAACCATGGGTGGACCCCGAAC	975
Qy	688	GGTGAGCGGGTCAGCG	AGAAACAAGACTCTGAAACGAGATCGCCGAGAGGGCGGCAACACC	747
Db	976	TCGCCGTGCTCTTG	GAGGACCCAGTCTTTGTGCTTGGCAAAAAGACAACGGAACC	1035
Qy	748	CTTGCTCAGGTTCTTAT	TGCTGGGTCTCGCGGTGCTACGTCGTTCTCCCCAGAGC	807
Db	1036	CCAGCCCTGATGCGCT	TACCACTGACGGTGGGGTTGTGTCCTGGCCCAAGC	1095
Qy	808	TCCAACCCCAAGCGCAT	TGAGTCCAACTTCAAGAGCAATTGAGTTC	852
Db	1096	TACAATGAGCAGGCGCAT	CAGACAGAAAGTGCAGGTGTTTGAATC	1140

RESULT 8

```

US-09-702-705-807
; Sequence 807, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 807
; LENGTH: 3829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-807

```

Db	677	ATTCAATTTCCAGTGTCTGTAAAGCCAGGTGAGGAAG-----TGATCCCAAAA	724
Qy	403	GGCAAAATAGCTATTCTCAAGGACCTGACCGAAGACCCGAGCCACATCGGCGCGCTATG	462
Db	725	GATGAAAATGGAAAAATACTATTTTGACACAGTGGATCTCTGTGCCACATGGGAGGCCATG	784
Qy	463	GAGAAGATTTATCAGATCGCAAGGGCCAGGTCATTGGTGTCTCCAACTGGACCATGGCC	522
Db	785	GAGAAGGTAAAGATCGAGGATTTGGCCAAAGTCCATCGGGGTGTCCAACTTCAACACACAGG	844
Qy	523	GACCTTTGAGAAGATGTCCTCAAGTTTCG-----CAAGGTTCATGCTCAGCGCAACCCAGATC	576
Db	845	CTGCTGGAGATGATCTCTACAGCCAGGGCTCAAGTACAAGCCTGTCTGCAACCCAGGTG	904
Qy	577	GAGATTCACCCCTTCTGCGCCCAAGAGGAGCTGGTGCAGTACTGTCTTCCAGAACAATT	636
Db	905	GAATGTATCTTACTTCAACACAGAGAAAACCTCTCGATTTCCTGCAAGTCAAAAGACATT	964
Qy	637	ATCCCGTGGCCCTACTCTCTCTGGGTGTCAGAACCCAGGTTCC-----CACCAACC	687
Db	965	GTTCCTGGTTCCTATAGTGTCTGGGATCCATCGAGAGAACCATGGGTGGACCCGAC	1024
Qy	688	GGTGAGGGGTGAGGAGAACAAAGATCTGAAAGAGATGCGCAGAGAGGGCGGCAACACC	747
Db	1025	TCCCCCGGTCTCTTGGAGGACCCAGTCTTTGTGCTTGGCAAAAAGCAACAGCGCAACC	1084
Qy	748	CTTGCTCAGGTTCTTATGCTGGGTCTGCGCGGTGGCTACGTGCTTCTCCCAAGAGC	807
Db	1085	CGAGCCCTGATTCGCTCGCTACCACTGACGCTGGGGTGTGTGCTCTGGCCAGAGC	1144
Qy	808	TCCAACCCCAAGCGCATTGAGTCCAACTTCAAGAGCATTTGAGCTC	852
Db	1145	TACAATGAGCAGCGCATCAGACAGAACTGTCAGGTGTTTGAAATC	1189

RESULT 9

US-09-736-457-807

Sequence 807, Application US/09736457

Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick

APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun

APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 21021.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 807

LENGTH: 3829

TYPE: DNA

ORGANISM: Homo sapiens

US-09-736-457-807

RESULT 9

```

US-09-736-457-807
Sequence 807, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 807
LENGTH: 3829
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-807

```

Db 916 GTTCGTGCCCTATAGTGCTCTGGGATCCCATCGAGAAGAACCATGGGTGGACCCGAAC 975

```
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-171

Query Match      10.9%; Score 106.6; DB 4; Length 1491;
Best Local Similarity 50.5%; Pred. No. 3.5e-21;
Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

QY 103 GCTGTCAACCACTGCCCTGAAGACCGGTTACCGTTCACCTGGACTGTCCTGGTACTACCTG 162
    |||||
Db 391 GCCGTCAAATGGCAATAGAGCCGGTTCACCATATTGATCTGCACATGTTTACAAT 450

QY 163 AACGAGGTGAGTGGTGGAGGTATCCGTGACTTCCTGAGAGAACCCCTCGGTGAAG 222
    |||||
Db 451 AATGAGGAGCAGGTTGGAGTGGCCATCCG---AAGCAAGATTGCAGATGGCAGTGTGAAG 507

QY 223 CQTGAGGACATCTTCGTCTGCACCAAGGTGTGGAAACCACTCCACCGTTATGAGGAGTC 282
    |||||
Db 508 AGAGAAGACATATTCTACCTCAAGCTTTGGAGCAATTCCTATCGACAGAGTTGGTC 567

QY 283 CTCTGGTCCATTGACACTCCCTGAAGCGTCTTGGACTGTACTAGTGTATGATGTTCTC 342
    |||||
Db 568 CGACAGCGCTTGGAAAGTCACTGAAAAATCTTCAATTGGACTATGTTGACCTCTATCTT 627

QY 343 GTTCACTGGCCATTGCTGCCAGAGAAATGGCCAGGTGAGCCCAAGATTGGCCCTGAC 402
    |||||
Db 628 ATTCAATTTCCAGTGTCTGTAAGCCAGGTGAGGAAG-----TGATCCCAAAA 675

QY 403 GGCAAAATAGCTCATTTCTCAAGGACCTGACCGAGAACCCCGAGCCCAATGCGCGCTATG 462
    |||||
Db 676 GATGAAATGGAAAAATACTATTGACACAGTGGATCTCTGTGCCACATGGAGGCCATG 735

QY 463 GAGAAGATTATGAGATCGMACCCAGGTCCATTGTGTCTCGAAGTGCACATGACCATGGCC 522
    |||||
Db 736 GAGAAGTGTAAAGATGCGAGGATTGGCCAAAGTCCATCGGGGTGTCCAACCTTCAACCAAGG 795

QY 523 GACCTTGAGAAGATGTCGAAGTTGCG-----CAAGGTCACTGCCTCAAGCAACAGATC 576
    |||||
Db 796 CTGCTGGAGATGATCTCTCAAGCCAGGGCTCAAGTACAGCCCTGTCTGCAACAGGTG 855

QY 577 GAGATTCAACCCCTTCTGCGCCAAACGAGGAGCTGGTGCAGTACTGTTCTCCAGAAACATT 636
    |||||
Db 856 GAATGTCACTCTTACTTCAACACAGAGAAATCTGCTGGATTTCTGCAAGTCAAAAGACATT 915

QY 637 ATGCCCGTGGCTACTCTCTCTGGCTCGAGAACCCAGGTTCC-----CACCACC 687
    |||||
Db 916 GTTCTGAGTATAGTGTCTCTGGATCCCATCGAAGAACCATGGGTGACCCGAAC 975

QY 688 GGTGAGCGGTCAGCGAGAACAAAGACTCTGAACGAGATCGCCGAGAGGGCGGCAACACC 747
    |||||
Db 976 TCCCCGGTGTCTTGGAGGACCCAGTCTTTGTGCTTGGCAAAAAGACAAAGCAAGCAACC 1035

QY 748 CTTGCTCAGGTTCTTATTGGCTGGGTCTGCGCGTGGCTAGTCTGTTCTCCCAAGAGC 807
    |||||
Db 1036 CCAGCCCTGATTGCCCTCGCTACAGCTGAGCGTGGGGTGTGGTCTCTGGCCAAAGAGC 1095

QY 808 TCCAAACCCCAAGCGATTGAGTCCAACTTCAAGAGCAATTGAGCTC 852
    |||||
Db 1096 TACAATGAGCAGCGCATCAGACAGAACGTCGAGGGTGTGTGATTTC 1140
```

RESULT 5

```
US-09-480-884A-171
; Sequence 171, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Panger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-171
```

```
Query Match      10.9%; Score 106.6; DB 4; Length 1491;
Best Local Similarity 50.5%; Pred. No. 3.5e-21;
Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

QY 103 GCTGTCAACCACTGCCCTGAAGACCGGTTACCGTTCACCTGGACTGTGCTGGTACTACCTG 162
    |||||
Db 391 GCCGTCAAATGGCAATAGAGCCGGTTCACCATATTGATCTGCACATGTTTACAAT 450

QY 163 AACGAGGTGAGTGGTGGAGGTATCCGTGACTTCCTGAGAGAACCCCTCGGTGAAG 222
    |||||
Db 451 AATGAGGAGCAGGTTGGACTGGCCATCCG---AAGCAAGATTGCAGATGGCAGTGTGAAG 507

QY 223 CGTGAAGACATCTTCGTCTGCACCAAGGTGTGAACCACTCCACCGTTATGAGGAGTC 282
    |||||
Db 508 AGAGAAGACATATTCTACACTTCAAGCTTTGGAGCAATTCCTATCGACAGATGGTTC 567

QY 283 CTCTGGTCCATTGACACTCCCTGAAGCGTCTTGGACTGTACTAGTGTATGATGTTCTC 342
    |||||
Db 568 CGACAGCGCTTGGAAAGTCACTGAAAAATCTTCAATTGGACTATGTTGACCTCTATCTT 627

QY 343 GTTCACTGGCCATTGCTGCCAGAGAAATGGCCAGGTGAGCCCAAGATTGGCCCTGAC 402
    |||||
Db 628 ATTCAATTTCCAGTGTCTGTAAGCCAGGTGAGGAAG-----TGATCCCAAAA 675

QY 403 GGCAAAATAGCTCATTTCTCAAGGACCTGACCGAGAACCCCGAGCCCAATGCGCGCTATG 462
    |||||
Db 676 GATGAAATGGAAAAATACTATTGACACAGTGGATCTCTGTGCCACATGGAGGCCATG 735

QY 463 GAGAAGATTATGAGATCGAAGGCCAGGTCCATTGTGTCTCCAACTGGACCATGGCC 522
    |||||
Db 736 GAGAAGTGTAAAGATGCGAGGATTGGCCAAAGTCCATCGGGGTGTCCAACCTTCAACCAAGG 795

QY 523 GACCTTGAGAAGATGTCGAAGTTGCG-----CAAGGTCACTGCCTCAAGCAACAGATC 576
    |||||
Db 796 CTGCTGGAGATGATCTCTCAAGCCAGGGCTCAAGTACAGCCCTGTCTGCAACAGGTG 855

QY 577 GAGATTCAACCCCTTCTGCGCCAAACGAGGAGCTGGTGCAGTACTGTTCTCCAGAAACATT 636
    |||||
Db 856 GAATGTCACTCTTACTTCAACACAGAGAAATCTGCTGGATTTCTGCAAGTCAAAAGACATT 915

QY 637 ATGCCCGTGGCTACTCTCTCTGGCTCGAGAACCCAGGTTCC-----CACCACC 687
    |||||
Db 916 GTTCTGAGTATAGTGTCTCTGGATCCCATCGAAGAACCATGGGTGACCCGAAC 975

QY 688 GGTGAGCGGTCAGCGAGAACAAAGACTCTGAACGAGATCGCCGAGAGGGCGGCAACACC 747
    |||||
Db 976 TCCCCGGTGTCTTGGAGGACCCAGTCTTTGTGCTTGGCAAAAAGACAAAGCAAGCAACC 1035

QY 748 CTTGCTCAGGTTCTTATTGGCTGGGTCTGCGCGTGGCTAGTCTGTTCTCCCAAGAGC 807
    |||||
Db 1036 CCAGCCCTGATTGCCCTCGCTACAGCTGAGCGTGGGGTGTGGTCTCTGGCCAAAGAGC 1095
```


Best Local Similarity 50.8%; Pred. No. 7e-25;
Matches 435; Conservative 0; Mismatches 388; Indels 33; Gaps 5;
QY 5 CTAACGGAAAGACTTTTCAATGACACACGGCGTCAAGATTCCTGGCGTCCGGCTTTGGTA 64
Db 12 CCATGGCAAGCGCTCTCTGCTTCAACAAACGGCGCAAGATCCCATCTGGGGTTGGTA 71
QY 65 CTTTCGCTAGTGAAGGTTCCRAAGGGCGAGACCTATCTGCTGTACCACTGCCCTGAAGA 124
Db 72 CCT-----GGAAGTCCCTCCAGGGCAGGTGACTGAGCGCGTGAAGTGCCATTGACG 125
QY 125 CCGGTTACCGTCACTTGACTGTGCTGGTACTACTCTGAACAGAGGTGAGTTGGTAGG 184
Db 126 TCGGTTACCGCACATCGACTGTGCCATGTGTACAGATGAGAATGAGGTGGGGTGG 185
QY 185 GTATCCGCTGACTTCTGTGAAGGAGACCCCTCGGTGAAGCGTGAAGACATCTTCTCTGCA 244
Db 186 CCAATTGAGAGAGCTGAGGAGAGGT---GGTGAAGCGTGAGGAGCTCTTCATCGTCA 242
QY 245 CCAAGGTGTGAACCACTCCACCGTTATGAGGAGCTCTCTGTGTCTTACGACACTCCC 304
Db 243 GCAAGCTGTGTGACGTACCATGAGAGGGCGCTGTGAAGAGGCTTCCAGAGACAC 302
QY 305 TGAAGCGTCTTGGACTTGACTGATGATGTTCTCTGTTCACTGGCCCATTTGCTGCCG 364
Db 303 TCAGCGACTGAAGCTGACTACCTGGACCTCTACTTTATTCACTGGCCGACTG----- 356
QY 365 AGAAGATGGCCAGGCTGAGCCCAAGATTGGCCCTGAGCGCAATATACGTCTTCTCAAGG 424
Db 357 -----GCTTTAAGCTGGAGAGATTTTCCCATTTGATGAGTCGGGCAATGGTTTC 410
QY 425 ACCTGACGAGAACCCGAGCCCATGCGCGCTATGAGAGATTTATGAGATCGCA 484
Db 411 CCAAGTGACACCAACATTTGACACAGCTGGCGGCCATGGAAGAGCTGGTGATGAAGGC 470
QY 485 AGGCCAGGTCCATTGGTGTCTCCAACTGGACCATTTGCCACTTGAAGATGTCCAA-- 542
Db 471 TGGTGAAGCTATTGGCATCTCCAACTTCAACCATCTCCAGGTGAGATGATCTTAACA 530
QY 543 ----GTTCCGCAAGTCTATGCTACGCCAACCAAGATCGAGATTAACCCCTTCTGCCCCA 598
Db 531 AACCTGGGTTGAAGTATGAAGCTGACGCTGATTAACCAAGATTGAGTGCACCATATCTCACTC 590
QY 599 AGGAGAGCTGTGAGTACTGCTTCTCCAGAACATTTATCCCGTGGCTACTCTCTC 658
Db 591 AGGAGAGTAAATCCAGTACTCCAGTCCAAAGGATCGTGGTACGCCCTTACAGCCCC 650
QY 659 TGGGCTCGCAGAACCAAGGTTCCCAACCCCGGTGAG-----CGGGTCAGCGAGAACCAAGA 712
Db 651 TGGGCTCTCTGACAGGCCCTGGGCGCAAGCCGAGGACCTTCTCTCTGGAGGATCCCA 710
QY 713 CTCTGAACGAGATCCCGAGAGGGCGGGAACACCCCTTGTCTCAGGTTCTTATTCGCTGGG 772
Db 711 GGATCAAGGCGATCGCAGCAAGCAATAAACTACAGCCAGGTCTGTATCCGGTTC 770
QY 773 GPTCTGGCGGTGGTACTGCTTCTCCCAAGAGCTCCAAACCCCAAGGCAATTGAGTCCA 832
Db 771 CCAATGACAGAGACTTGTGTGTATCCCAAGTCTGTGACACCAAGAGCAATGCTGAGA 830
QY 833 ACTTCAAGAGCATTTGA 848
Db 831 ACTTTAAGTCTTTGA 846

RESULT 2
US-08-801-344-3
; Sequence 3, Application US/08801344
; Patent No. 6087140
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Altaras, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL

; TITLE OF INVENTION: FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,344
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat lens aldose reductase gene
; US-08-801-344-3

Query Match 11.6%; Score 113; DB 3; Length 1337;
Best Local Similarity 50.7%; Pred. No. 4.5e-23;
Matches 424; Conservative 0; Mismatches 380; Indels 33; Gaps 5;
QY 24 ATTGAGCAACGGCGTCAAGATTCTTGGCGCTTGTGTGCTTGGTACCTTCGCTAGTGAAGGTTC 83
Db 53 ACTCAACACGGCACCAAGATGCCACCTGGTCTGGCACCT-----GGAAGTCTCC 106
QY 84 CAAGGCGGACCTATACCTGTGTACCACTGCCCTGAAAGACCGGTACCGTCACTTGA 143
Db 107 TCCTGGCCAGGTGACCGAGGCTGTGAAGGTGTGTATCGACATGGGGTATCGCCACATTGA 166
QY 144 CTGTGCTCGTACTACTGAAAGGGTGTGTGTGAGGGTATCCGTGACTTCTCTGAA 203
Db 167 CTGCGCCGAGGTGTACAGAAATGAGAGAGGTGGGGTGGCCCTCCAGGAGAGCTCAA 226
QY 204 GGAGAACCCCTCGGTGAAGCGTGAAGACATCTTGTCTGACCAAGGTGTGAAACCACT 263
Db 227 GGAGCAGGT---GGTGAAGCGGAGGATCTCTTCTTGTGTGAGAGCTGTGGTCAAGCTT 283
QY 264 CCACCGTTATGAGGACGCTCTCTGTGTCCATTGACGACTCCCTGAAAGCGTCTTGGATTGA 323
Db 284 CCAGACGAGAGATGTTGAAGGGGCGCTTGCAGAGAGACGCTGAGCGACCTGCAGCTGA 343
QY 324 CTACGTTGATATGTTCTCGTTCACTGGGCCCTTGTGCGGAGAGAAATGCGCAGGTTGA 383
Db 344 CTACCTGGACCTCTACCTTTTACTTGGCCCACTG-----GCTTCAAGGCTCG 391
QY 384 GCCCAAGATTGGCCCTGACGGCAATACGTAATCTTCAAGGACCTGACCGAGAACCCCA 443
Db 392 GCCTGACTATTTCCCGCTGGATGTCATCGGAAACCGTATTCTTAGTGACACCGATTTGT 451
QY 444 GCCCAGATGGCGCTATGAGAGATTTATGAGGATCGCAAGGCGAGGTCCATTGGTGT 503
Db 452 GGCACCTTGGACGGCTATGAGCAACTAGTGTGATGAGGTTTGGTAAAGCAATCGAGT 511

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2004, 01:05:39 ; Search time 76 Seconds
(without alignments)
7141.342 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978
Sequence: 1 atgtctaaggaagacttt.....ccaagaacctgtctgcgtga 978

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/6C COMB.seq: *
6: /cgn2_6/ptodata/2/ina/6D COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119.2	12.2	1335	4	US-09-023-655-1010
2	113	11.6	1337	3	US-08-801-344-3
3	113	11.6	1337	4	US-09-498-599-3
4	106.6	10.9	1491	4	US-09-643-597-171
5	106.6	10.9	1491	4	US-09-480-884A-171
6	106.6	10.9	1491	4	US-09-542-615A-171
7	106.6	10.9	1491	4	US-09-606-421B-171
8	106.6	10.9	3829	4	US-09-702-705-807
9	106.6	10.9	3829	4	US-09-736-457-807
10	106.6	10.9	3829	4	US-09-614-124B-807
11	106.6	10.9	3829	4	US-09-671-325-807
12	106.6	10.9	3829	4	US-09-589-184-807
13	106.6	10.9	3829	4	US-09-702-705-784
14	106.6	10.9	3829	4	US-09-736-457-784
15	106.6	10.9	3829	4	US-09-614-124B-784
16	106.6	10.9	3829	4	US-09-671-325-784
17	106.6	10.9	3829	4	US-09-589-184-784
18	106.6	10.9	3829	4	US-09-702-705-95
19	103.8	10.6	1265	4	US-09-736-457-95
20	103.8	10.6	1265	4	US-09-614-124B-95
21	103.8	10.6	1265	4	US-09-671-325-95
22	103.8	10.6	1265	4	US-09-589-184-95
23	103	10.5	1199	4	US-09-519-172-80
24	103	10.5	1199	4	US-09-976-594-1079
25	101	10.3	1683	3	US-09-347-803-11
26	99.2	10.1	1316	4	US-09-702-705-323
27	99.2	10.1	1316	4	US-09-736-457-323

28	99.2	10.1	1316	4	US-09-614-124B-323	Sequence 323, Appl
29	99.2	10.1	1316	4	US-09-671-325-323	Sequence 323, Appl
30	99.2	10.1	1316	4	US-09-589-184-323	Sequence 323, Appl
31	98.8	10.1	1207	3	US-09-166-412-1	Sequence 1, Appl
32	98.8	10.1	1207	4	US-08-731-320B-1	Sequence 1, Appl
33	98.8	10.1	1207	4	US-09-166-438-1	Sequence 1, Appl
34	95	9.7	1206	3	US-08-532-896-1	Sequence 1, Appl
35	83.6	8.5	1132	4	US-09-626-002-18	Sequence 18, Appl
36	79.2	8.1	1233	3	US-08-853-839-1	Sequence 3, Appl
37	78.2	8.0	1259	3	US-09-166-412-3	Sequence 3, Appl
38	78.2	8.0	1259	4	US-08-731-320B-3	Sequence 3, Appl
39	78.2	8.0	1259	4	US-09-166-438-3	Sequence 3, Appl
40	77	7.9	954	2	US-08-336-198C-2	Sequence 2, Appl
41	77	7.9	954	4	US-09-184-965-2	Sequence 2, Appl
42	75.2	7.7	1073	3	US-09-347-803-5	Sequence 5, Appl
43	72.8	7.4	530	4	US-09-702-705-87	Sequence 87, Appl
44	72.8	7.4	530	4	US-09-736-457-87	Sequence 87, Appl
45	72.8	7.4	530	4	US-09-614-124B-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1

US-09-023-655-1010
; Sequence 1010, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1010:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1335 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g179035
US-09-023-655-1010

Query Match

12.2%; Score 119.2; DB 4; Length 1335;

Db 532 GAGGAAGCTTCGGCGAAGCTTCCTCGGTAAACAGATCGAAGCTCAACCCCTTGTGCAA 591
QY 598 AACGAGAGCTGGTGCAGTACTGCTTCTCCAGAACATTAATGCGGTGGCTACTCTCCT 657
Db 592 CAACAGAGTGGTGTCTCACCATAAGAGCAAGAACATGTGTCATCTGCTTACAGTCCC 651
QY 658 CTGGGCTCGCAAGAACAGGTTCCACCACCGGTGAGCGGTGACGAGAAACAAGACTCTG 717
Db 652 TTGGGCAACAATGTGCTGTGTAAACACCTC-----TGACTGAGATCCGGTATT 702
QY 718 AACGAGATCGCGAGAGGGGGGCAACACCTCTCTCAGGTTCTTATGCTGGGTCTG 777
Db 703 GTGATGCTGTAGAGCTCTGAACCATCTCTCTGCTGTGCTCATGCTTGGGGTATT 762
QY 778 CGCGTGTGCTACGCTCTCTCCCAAGAGCTCCAAACCCCAAGCGCAATTGATCCAACTTC 837
Db 763 CAACGGGTACAGGCTCTTGGTCAAGTCAAGTACAGCTTACACCTCACGGATTAGAGTAACTTT 822
QY 838 AAGAGCATGAGCTCTCGGATGCGGACTTTGAAGCCATCAATGCGGTGGCAAGG 892
Db 823 GAACAGATCACTCTGTCTGATGAGGAATTCACAGGGTTACCAACCTCATCAAGG 877

RESULT 13
US-10-425-114-17144
; Sequence 17144, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 17144
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3068-021-G2_FLI
US-10-425-114-17144

Query Match 13.4%; Score 130.8; DB 13; Length 1289;
Best Local Similarity 52.4%; Pred. No. 3.4e-31;
Matches 416; Conservative 0; Mismatches 357; Indels 21; Gaps 5;

QY 107 TCACACTGCTCGAAGACGGTTACCGTCACTTGGACTGTGCTGCTACTACTGAAAG 166
Db 185 TCCACAGCGCTCCGGTCCGCTGCGTACCGCCACCTGGACTGCGCGCTGACTACAGAACG 244
QY 167 AGGTGAGTTGGTCAAGGCTTCGGTCACTTCTCCAGAGAACCCCTCGGTGAAGCGTG 226
Db 245 AAGCTGAGTTGGTCAAGCGCTCGCAGAGGCAATCCAGACCGGACTC---GTCAAGCGG 301
QY 227 AGGACATCTTCTGTGCAACCAAGGTGGAAACCACTCCACCGTTATGAGACGCTCTCT 286
Db 302 AGGACCTTCTCATCACACCAAGCTGTGGAACACTCAGACC-----ATGGCCATGTGTTG 355
QY 287 GGTCCATTGACGACTCCCTGAAGCGTCTTGACTTTGACTAGTTGATGTTCCTCGTTC 346
Db 356 AAGCTGCAAGGACAGCTTGAAAGCTGACGTAGACTATCTCGACTCTTACCTCATCC 415
QY 347 ACTGCCCCATTGCTGCGAGAAAGAAATGCCAGGTGAGCCCAAGATTGGCCCTGACGGCA 406
Db 416 ATTTCCAGTAGCCACTCGGCACACCGGAGTCGGCAGCACTTCTAGCGCTCTGGGTGACG 475
QY 407 AATAGTCACTTCAAGGACCTGACCGGAGAACCCCGGACCCACATGCGCGCTATGAGA 466

Db 476 ATGGCGTGTGACATCGACACCACTATCTCCTCGAAAGACATGGCACCGATGTAAG 535
QY 467 AGATTATAGAGATCGAAGGCCAGGTCCATTGGTGTCTCCAACTGGACCAATGCGGACC 526
Db 536 AGCTTGTGTTCCATGGGGTGTGCGCAGCATTTGGAATCAGCAACTACGACATCTTCTCA 595
QY 527 TTGAGAAGATGTCCTAAGTTTCGCAAGGTCATGCCCTCAGCGCAACCAAGATCGAGATTCA 586
Db 596 CCAGAGACTGCTGCGCTACGCCAAGATAAAGCGCGGTGAACAGATCGAGACGAC 655
QY 597 CTTCTCTGCCAAACAGAGAGCTGGTGCAGTACTGC---TTCTCAAGAACATATTGCGCG 643
Db 656 CTTACTTCAGCGCGACTCTCTTGTCAAGTCTGCGCAGAACACCGGATCTGCGTGACG 715
QY 644 TGGCTACTCTCTCTGCGCTCGCAGAACCAAGTTCCACACCGGTGAGCGGGTCAAGG 703
Db 716 CGCACACCGCTGGTGGTCCACCGCAACCGCGAGTGGTTCGGGAGGGTCTCGTGGC 775
QY 704 AGAACAAAGACTCTGAACGAGA---TCGCGAGAAAGGCGGCAACCCCTTGTCTAGGTT 760
Db 776 TGGACACCTCTCATCAAGAGCTTTGGCTGAGAAGTACGGCAAGACGCGCGCGAGCTCG 835
QY 761 TTATTGCTGGGTCTGCGCGTGGCTAGTGTCTTCCCAAGAGCTCCACCCCAAGC 820
Db 836 TCTCTGCTGGGTCTGACGCGGACACGTTGGTGTATCCCAAGACCTCCAAAGCGGAGA 895
QY 821 GCATTGAGTCCAACCT-----TCAAGAGCATTTGAGCTCTCCGATCCGACTTTGAAGCCA 874
Db 896 GGCTGCAGGAGAACTTTGACGCTTTCGGTTTTTGACATCTCCGCGGAGGACATGGAGAGA 955
QY 875 TCAATGCCGTGTC 888
Db 956 TGAAGCGCGTGCAC 969

RESULT 14
US-10-425-114-25069
; Sequence 25069, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25069
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-245-C12_FLI
US-10-425-114-25069

Query Match 13.2%; Score 129.4; DB 13; Length 1289;
Best Local Similarity 51.8%; Pred. No. 9.5e-31;
Matches 398; Conservative 0; Mismatches 356; Indels 15; Gaps 4;

QY 107 TCACACTGCTCGAAGACCGGTACCGTCACTTGGACTGTGCTGCTACTACTGAAAG 166
Db 180 TCACACAGCGTTCGGCTCGGCTACCGCCACTCGACTCGCGCTGACTTACCAAGAACG 239
QY 167 AGGTGAGTTGGTGAAGGTTATCCGTGATCTTCTTGAAGAGAACCCCTCGGTGAAGCGTG 226
Db 240 AAGCTGAGTTGGTGAAGCGCTCGCAGAGGCAATCCAGACCGGACTC---GTCAAGCGG 296

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Trichosporonoides megachiliensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(987)
; OTHER INFORMATION:
US-10-040-416-5

Query Match      14.7%; Score 144.2; DB 14; Length 1121;
Best Local Similarity 50.9%; Pred. No. 1.6e-35;
Matches 456; Conservative 0; Mismatches 418; Indels 21; Gaps 4;

QY 1 ATGCTTAACGGAAGACTTTCACATTTAGACACGCGCTCAAGATTCCTGCGGTTCGGTTT 60
DB 1 ATGCTTACAAACAAGACATCCCTCTCAACGACGCTAACTCCATCCCTGCGCTTGGCTAC 60
QY 61 GGTACCTTCGCTAGTGAAGTTTCCAAAGGGGAGACCTTATCTGCTGTCACCACTGCCCTG 120
DB 61 GGTACCTT-----GGCAAGCAGAACCTTGTGAGTCCGCTGAGTGTCAAGCTCGCTGTC 114
QY 121 AAGACCGGTATACCTGCTACTTGGACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 115 AAGCTGGCTACCGTCACTTGGACTTGGCCAAAGTTTACCAGAACCAACCGAGATTGGC 174
QY 181 GAGGTATCCGTGACTTCCCTGAGGAGAACCCCTCGGTGAGCGTGAAGACATCTTCGTC 240
DB 175 CAAGCTCTCAAGAACCTGTTTGAAGGGTGT---GTCAGCGTGAAGACCTTTTCATC 231
QY 241 TGCACCAAGGTGTGGAACCCACCTCCACCGTTATAGAGGAGCTCTCTGCTCATTGACAC 300
DB 232 ACTTCCAAAGCTTTGGAACCAACCGCCACCGCTCCTGAGACAGCTTGAGCCCTGCGCTC 291
QY 301 TCCCTGAAGCGTCTGGACTTGACTACCTGATATGTTCTCTGTTCACTGGCCCACTTGC 360
DB 292 ACTCTTAAGAGTTCGGTCTATCTCTATTTGACCTGTACTTGAATTCATCTGCTGCTTGC 351
QY 361 GCCGAGAAGAATGGCCAGGGTGA---GCCAAGATTGGCCCTGACGGCAATACCTCAT 417
DB 352 TTCAAGTTCACTACTCTCCGATGAAGTGTCTCTGCTGACCTTACCAACAGGACCTTGCC 411
QY 418 CTCAGGACCTGACAGAACCCCGAGCCCATGCGCGCTATGAGAGATTTATGAG 477
DB 412 TACATTGAGATTCCGTCATAATGTCGCACACCTTGAAGCGGCTGCTGCGCTCGAATAAG 471
QY 478 GATCGCAAGCCAGCTCCATTGGTGTCTCCAACTGGACCACTTGGCCAGCTTTGAGAGATG 537
DB 472 ACGGGTAAGACCAAGTCGGTTGGTGTTCGAACTTCAGACATCTGCTCTGGTCTGCTTGC 531
QY 538 TCCAAAGTTCCGAAGTTCGCTGATGAGGAAATTCACACGGGTACCAACCTCATCAAGG 877
DB 523 AAGAGCATGTGCTGCTGATGAGGAAATTCACACGGGTACCAACCTCATCAAGG 877

RESULT 12
US-10-040-416-3
; Sequence 3, Application US/10040416
; Publication No. US20020160480A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, TETUYA
; APPLICANT: KASUMI, TAKAFUMI
; APPLICANT: ASABA, EIJI
; TITLE OF INVENTION: ERYTHROSE REDUCTASE, ITS cDNA AND CELL WHICH THE cDNA EXPRESS
; FILE REFERENCE: 217969USOXIP
; CURRENT APPLICATION NUMBER: US/10/040,416
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 09/800,487
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: JP 2001-001294
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Trichosporonoides megachiliensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(987)
; OTHER INFORMATION:
US-10-040-416-3

Query Match      14.6%; Score 142.6; DB 14; Length 1077;
Best Local Similarity 50.8%; Pred. No. 5.1e-35;
Matches 455; Conservative 0; Mismatches 419; Indels 21; Gaps 4;

QY 1 ATGCTTAACGGAAGACTTTCACATTTAGACACGCGCTCAAGATTCCTGCGGTTCGGTTT 60
DB 1 ATGCTTACAAACAAGACATCCCTCTCAACGACGCTAACTCCATTCCTGCTTCTTGGGTAC 60
QY 61 GGTACCTTCGCTAGTGAAGTTTCCAAAGGGGAGACCTTATCTGCTGTGCACCACTGCCCTG 120
DB 61 GGTACCTT-----GGCAAGCAGAACCTTGTGAGTCCGCTGAGGTTGTCAAGCTCGCTGTT 114
QY 121 AAGACCGGTATACCTGCTACTTGGACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 115 AAGGTGGCTACCGTCACTTGGACTTGGCCAAAGTTTACCAGAACCAACCGAGATTGGC 174
QY 181 GAGGTATCCGTGACTTCCCTGAGGAGAACCCCTCGGTGAGCGTGAAGACATCTTCGTC 240
DB 175 CAAGCTCTCAAGAACCTGTTTGAAGGGTGT---GTCAGCGTGAAGACCTTTTCATC 231
QY 241 TGCACCAAGGTGTGGAACCCACCTCCACCGTTATAGAGGAGCTCTCTGCTCATTGACAC 300
DB 232 ACTTCCAAAGCTTTGGAACCAACCGCCACCGCTCCTGAGACAGCTTGAGCCCTGCGCTC 291
QY 301 TCCCTGAAGCGTCTGGACTTGACTACCTGATATGTTCTCTGTTCACTGGCCCACTTGC 360
DB 292 ACTCTTAAGAGTTCGGTCTATCTCTATTTGACCTGTACTTGAATTCATCTGCTGCTTGC 351
QY 361 GCCGAGAAGAATGGCCAGGGTGA---GCCAAGATTGGCCCTGACGGCAATACCTCAT 417
DB 352 TTCAAGTTCACTACTCTCCGATGAAGTGTCTCTGCTGACCTTACCAACAGGACCTTGCC 411
QY 418 CTCAGGACCTGACAGAACCCCGAGCCCATGCGCGCTATGAGAGATTTATGAG 477
DB 412 TACATTGAGATTCCGTCATAATGTCGCACACCTTGAAGCGGCTGCTGCGCTCGAATAAG 471
QY 478 GATCGCAAGCCAGCTCCATTGGTGTCTCCAACTGGACCACTTGGCCAGCTTTGAGAGATG 537
DB 472 ACGGGTAAGACCAAGTCGGTTGGTGTTCGAACTTCAGACATCTGCTCTGGTCTGCTTGC 531
QY 538 TCCAAAGTTCCGAAGTTCGCTGATGAGGAAATTCACACGGGTACCAACCTCATCAAGG 877
DB 523 AAGAGCATGTGCTGCTGATGAGGAAATTCACACGGGTACCAACCTCATCAAGG 877
```

```
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: JP 2000-372704
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: JP 2001-006144
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: JP 2001-026594
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: JP 2001-175175
; PRIOR FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 18
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (14)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-10-004-115A-18

Query Match 27.1%; Score 265; DB 15; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.4e-74;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTAACGGAAGACTTTTCAATTTGAGCAACGGCGTCAAGATTCTTGGCGTGGCTTT 60
DB 67 ATGCTAACGGAAGACTTTTCAATTTGAGCAACGGCGTCAAGATTCTTGGCGTGGCTTT 126
QY 61 GGTACTTTCGGTGTAGTGAAGTTTCAAGGGCGAGACCTTACTGTGTCAACACTTGCCTG 120
DB 127 GGTACTTTCGGTGTAGTGAAGTTTCAAGGGCGAGACCTTACTGTGTCAACACTTGCCTG 186
QY 121 AAGCCGGTTACCGTCACTTGGACTGTGCTGTGTTACTTGAACGAGGGTGGTGGT 180
DB 187 AAGCCGGTTACCGTCACTTGGACTGTGCTGTGTTACTTGAACGAGGGTGGTGGT 246
QY 181 GAGGTATCCGTGACTTCTCAAGAGAACCCCTCGGTGAGCGTGAGGACATCTTCGTC 240
DB 247 GAGGTATCCGTGACTTCTCAAGAGAACCCCTCGGTGAGCGTGAGGACATCTTCGTC 306
QY 241 TGCACCAAGGTGTGGAACCACTCC 265
DB 307 TGCACCAAGGTGTGGAACCACTCC 331

RESULT 10
US-09-734-237B-31
; Sequence 31, Application US/09734237B
; Publication No. US20030064432A1
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Sporidiobolus salmonicolor
; US-09-734-237B-31

Query Match 20.4%; Score 199.6; DB 13; Length 972;
Best Local Similarity 55.5%; Pred. No. 2.3e-53;
Matches 463; Conservative 0; Mismatches 344; Indels 27; Gaps 3;

QY 87 GGGCGAGACTTACTGTGTGTCACCACTGCCCTGAGAGACCGGTTACCGTCACTTGGACTG 146
```

```
DB 75 GGGCGAGGTGGCGCCAGGGCGTCAAGGTGCCATCGAGACTGGATACCGTCACTCGACT 134
QY 147 TGCCTGTGTTACTTACCTGACGAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 206
DB 135 TCCCAAGGTCTTACTCGAACCAACCTGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 179
QY 207 GAACCCCTCGGTGAAGCGTGAAGGACATCTTCTGTCACCAAGGTGTGGAACCACTCCA 266
DB 180 GGAGGCTGGCTCAAGCGGAGGACCTTTTATCACTCTGAGCTCTGGAACAATCGCA 239
QY 267 CCGTTATGAGGACGTCTTCTGTCATTTGACGACTTCCCTGAAAGCGTCTTGGACTTGA 326
DB 240 CCGCCCGGAGCAGGTGAGCGCTTCCCTTACGACACACCTCTCAAGGAGCTCGGCTCGAG 299
QY 327 CGTTGATATGTTCTCTGTTCACTGGCCCATTTCTGCCGAGAGATGCGCAGGGTGAGCC 386
DB 300 CTTCGACCTTTTACCTTCACTTGGCCCGTGGCTTCCGCCCGGAGGGGAGCATCACCA 359
QY 387 CAAGATTGGCCCTGACGCAAAATACGTCTTCTCAAG---GACCTGACCGAAGACCCCA 443
DB 360 GAACCTCTTCCCGAAGGCCAACGACAGGAGGTCAAGCTCGACCTCGAGGTGAGGTCTGT 419
QY 444 GCCCATGCGCGCTATGAGAGATTTATGAGGATCGCAGGCGAGGTCCATTTGGTGT 503
DB 420 CGACAGCTGGAAGGGGATGTTCAAGCTTCTCGACACTGSCAAGGTCAAGGGCGATCGGG 479
QY 504 CTCCAACTGGACCATTTGCCGACCTTGAGAAAGATGTCGAAGTTTCGCAAGGTTCATGCC 563
DB 480 TTCCAACTTCGAGCGGAGATGGTGCAGCCATCATCGAGGCTACCGGCGTCAACCCCTC 539
QY 564 CGCAACAGAGATCGAGATTCACCCCTTCTGCTGCCAACGAGGAGCTGGTGCAGTACTGTT 623
DB 540 CGTCAACAGATCGAGCGTCAACCTTCTCTTCAGCCCGGAGCTCATGCCCAACCAAA 599
QY 624 CTCCAAGAACATTTAGCCGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
DB 600 GSCCAAGACATTCATATTACCGCATCTCTCTCTCGTAAACACACCGTGGCGGCGCC 659
QY 684 CACCGGTGAGCGGTGAGCGAGAAACAAGACTCTGAACGAGATCGCGGAGAGGGGGGCA 743
DB 660 TCTTCTTGTCCAGCACCCCGAGATCAAGCGC-----ATCGCGGAGAAACGGGTG 710
QY 744 CACCTTGTCTCAGGTCTTTATTTGCTGGGCTCTGCGCGCTGCTGCTGCTGCTGCTGCT 803
DB 711 CACCGCGCTCAGTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
QY 804 GAGCTCCAAACCCCAAGCGCATTTGAGTCCAACTTCAAGAGCATTTGAGTCTCCGATGCC 863
DB 771 GTCGGTCAACCCCTCCCGCATTTGGCGAGAACTTCAAGCAGGTCTGCTCTCGCAGGAG 830
QY 864 CTTTGAAGCATCAATGCCGTTGCCAAGGTCGTCACTTCCGTTTTCGTCAACAT 917
DB 831 CGTCGATGCCGTGAGCAAGCTCGGCGAGGGTTCGGGCGCGAGCGCTCAACAT 884
```

```
RESULT 11
US-10-040-416-5
; Sequence 5, Application US/10040416
; Publication No. US20020160480A1
; GENERAL INFORMATION:
; APPLICANT: OOKURA, TETUYA
; APPLICANT: KASUMI, TAKAFUMI
; APPLICANT: ASABA, EIJI
; TITLE OF INVENTION: ERYTHROSE REDUCTASE, ITS CDNA AND CELL WHICH THE CDNA EXPRESS
; FILE REFERENCE: 217969USOXCIIP
; CURRENT APPLICATION NUMBER: US/10/040,416
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 09/800,487
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: JP 2001-001294
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 10
```

FEATURE:
 NAME/KEY: modified_base
 LOCATION: (717)
 OTHER INFORMATION: a, c, t, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (723)
 OTHER INFORMATION: a, c, t, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (728)
 OTHER INFORMATION: a, c, t, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (734)
 OTHER INFORMATION: a, c, t, g, other or unknown

US-10-004-115A-19

Query Match 42.1%; Score 412; DB 15; Length 743;
 Best Local Similarity 100.0%; Pred. No. 9.9e-122;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 CAACAGATCAGATTCACCCCTTCTGCGCCACGAGGAGCTGGTGCAGTACTGCTTCTC 626
 DB 1 CAACAGATCAGATTCACCCCTTCTGCGCCACGAGGAGCTGGTGCAGTACTGCTTCTC 60

QY 627 CAAGAACATATTGCGCGTGGCTACTCTCTCTGCGGCTCGCAGAACAGGTTCCACAC 686
 DB 61 CAAGAACATATTGCGCGTGGCTACTCTCTCTGCGGCTCGCAGAACAGGTTCCACAC 120

QY 687 CGGTGAGCGGTGAGGAGAACAGACTCTGACGAGATGCCGAGAAAGGCGGCGACAC 746
 DB 121 CGGTGAGCGGTGAGGAGAACAGACTCTGACGAGATGCCGAGAAAGGCGGCGACAC 180

QY 747 CTTGCTCAGGTTCTATTGCTGCGGTCTGCGCGGCTGCTGCTGCTTCTTCCCAAGAG 806
 DB 181 CTTGCTCAGGTTCTATTGCTGCGGTCTGCGCGGCTGCTGCTGCTTCTTCCCAAGAG 240

QY 807 CTCCAACCCCAAGGCGATTGAGTCCAACTTCAAGAGATTGAGTCTCCGATGCCGACTT 866
 DB 241 CTCCAACCCCAAGGCGATTGAGTCCAACTTCAAGAGATTGAGTCTCCGATGCCGACTT 300

QY 867 TGAAGCCATCAATCCGTTGCGAGGTCGTCACCTCCGTTTCTGTCACATGAGGATAC 926
 DB 301 TGAAGCCATCAATCCGTTGCGAGGTCGTCACCTCCGTTTCTGTCACATGAGGATAC 360

QY 927 TTTGGATATGATGCTGCGCGCGAGGAGACCGCCAAAGAACCTGTCGCTGA 978
 DB 361 TTTGGATATGATGCTGCGCGCGAGGAGACCGCCAAAGAACCTGTCGCTGA 412

RESULT 8
 US-10-369-493-27083
 ; Sequence 27083, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10152052B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIORITY FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 27083
 ; LENGTH: 765
 ; TYPE: DNA
 ; ORGANISM: Neurospora crassa

US-10-369-493-27083

Query Match 41.7%; Score 408.2; DB 16; Length 765;
 Best Local Similarity 71.8%; Pred. No. 1.7e-120;
 Matches 549; Conservative 0; Mismatches 213; Indels 3; Gaps 1;

QY 46 CCTGGCGTGGCTTTGGTACCTTCCTAGTAGAAGTTTCAAGGGCGAGCACTATCTCT 105
 DB 1 CCGGCGTGGATTTCGGCACCTTTGCCAACGAGGTTGCCAACAGGGCGAGCACTACGCGCC 60

QY 106 GTCACCACTGCCCTGAAGACCGTTACCGTCACTTGGACTGTGCTGCTACTACTCTGAAC 165
 DB 61 GTCACCAAGGCGCTCGAGTTGGCTACCGTCACTTGGACTGTGCTGCTACTACTCTGAAC 120

QY 166 GAGGCTGAGTTGGTGAAGGTATCCGTGACTTCTTGAAGGAGAACCCCTCGTGAAGCGT 225
 DB 121 GAGGAGAGGTGCGGCGAGCGCTCCCGAGTTTCTTTGAGAACCAACAGGACGCTCAAGCGC 180

QY 226 GAGGACATCTTCTGTGACCAAGGTGTGAACCACTTCCACCGTTATGAGGACGCTCTC 285
 DB 181 GAGGACATCTTCTGTGACCAAGGTGTGAACCACTTCCACCGTTATGAGGACGCTCTC 240

QY 286 TGTGCTCATTTGAGACTTCCCTGAAGCGTCTTGGACTTGAATGATGTTCTCTCTT 345
 DB 241 TGTGCTCATTTGAGACTTCCCTGAAGCGTCTTGGACTTGAATGATGTTCTCTCTT 300

QY 346 CACTGCGCCCATTTGCTGCGGAGAG--AATGGCCAGGTGAGCCCAAGATTGGCCCTGAC 402
 DB 301 CACTGCGCCCATTTGCTGCGGAGAGAGCGAGGCCACCAACATGCCCAAGATCGGCCCGCAT 360

QY 403 GGCACATAGCTATTCTCAAGGACCTGACGAGACCGGAGACCGGCGCCACATGCGGCGCTATG 462
 DB 361 GGCACATAGCTATTCTCAAGGAGGCTCACCGAGAACCGGAGACCGGCGCGGCGCTATG 420

QY 463 GAGAAGATTATGAGGATCGCAAGCGCCAGGTCTCATTTGGTGTCTTCCAACTGGACCATTCGC 522
 DB 421 GAGGATCTCGTGATGCGGCGAAGACGCGTCTCATTTGGTGTCTTCCAACTGGACCATTCGC 480

QY 523 GACCTTGAGAAGATGTCGAAGTTCGCAAGGTCTGCTCATCGCCCAACAGATCGAGATT 582
 DB 481 GGCCTCGAGAAGCTCTCAAGTTTCTCGCATCAAGCGCGCCGTCACACAGATTGAATC 540

QY 583 CACCCCTTCTGCGCCCAAGGAGGCTGTGCACTACTCTTCTCCAGAACTATGCGCC 642
 DB 541 CACCCCTTCTGCGCCCAAGGAGGCTGTGCACTACTCTTCTCCAGAACTATGCGCC 600

QY 643 GTGGCTACTTCTCTTCTGGGTTCGAGAACGAGGTTCCTCAACCGGTGAGCGGTCAGC 702
 DB 601 ACTGCTACTTCTCTCTGCGTTCGAGAACGAGGTTCCTCAACCGGTGAGCGGTCAGC 660

QY 703 GAGAAGACTTCTGAAGATCGCGAGGCGGAGGCGGCGGCGGCGGCGGCTCTT 762
 DB 661 GATGACCCCACTCTCAAGGCGGTTCTGAGCGGCGGAGACCACTTGGCCAGGTCTC 720

QY 763 ATTGCTGGGCTCTGCGGCTGCTGCTGCTGCTGCTGCTTCTTCCCAAGAGC 807
 DB 721 CTCGCTGGGCTCTCGCAGAGGCTACGTCGTCCTCCCAAGAGC 765

RESULT 9
 US-10-004-115A-18
 ; Sequence 18, Application US/10004115A
 ; Publication No. US20030134402A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ASAKO, HIROYUKI
 ; APPLICANT: NATSUMURA, KENJI
 ; APPLICANT: SHIMIZU, MASATOSHI
 ; APPLICANT: ITO, NOBUJA
 ; APPLICANT: WAKITA, RYUHEI
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
 ; TITLE OF INVENTION: 4-HALO-3-HYDROXYBUTANOATE
 ; FILE REFERENCE: 7372-72249
 ; CURRENT APPLICATION NUMBER: US/10/004,115A

Qy 605 AGCTGGTGCAGTACTGTTCTCCAGAACATTATGCCGGCGGCTACTCTCCTCTGG 661
Db 361 AGCTGGTGCAGTACTGTTCTCCAGAACATTATGCCGGCGGCTACTCTCCTCTGG 417

RESULT 7

US-10-004-115A-19
; Sequence 19, Application US/10004115A
; Publication No. US20030134402A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: MATSUMURA, KENJI
; APPLICANT: SHIMIZU, MASATOSHI
; APPLICANT: ITO, NOBUYA
; APPLICANT: WAKITA, RYUHEI
; TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
; FILE REFERENCE: 7372-72249
; CURRENT APPLICATION NUMBER: US/10/004,115A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: JP 2000-372704
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: JP 2001-006144
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: JP 2001-026594
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: JP 2001-175175
; PRIOR FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (434)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (440)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (443)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (448)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (458)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (467)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (475)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (506)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (510)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (517)
; OTHER INFORMATION: a, c, t, g, other or unknown
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (522)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (533)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (535)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (583)..(584)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (607)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (616)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (634)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (639)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (648)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (651)..(652_)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (659)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (662)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (664)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (670)..(671)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (690)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (693)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (695)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (702)
; OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE:
NAME/KEY: modified_base
LOCATION: (463)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (485)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (489)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (491)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (520)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (524)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (593)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (602)
OTHER INFORMATION: a, c, t, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (682)
OTHER INFORMATION: a, c, t, g, other or unknown
US-10-004-115A-15

Query Match 60.2%; Score 588.6; DB 15; Length 697;
Best Local Similarity 94.4%; Pred. No. 1.5e-178;
Matches 611; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
QY 4 TCTAACGGAAGACTTTCACATTCAGCAACGGCGTCAAGATTCCTGCGCTCGCTTGGT 63
DB 52 TCCAAACGGAANACTTTCACACTGAGCAACGGCGTCAAAATTCCTGGCGTGGCTTGGT 111
QY 64 ACCTTCGCTAGTGAAGGTTCCAAAGGGGAGACCTTATCTGTGTCAACACTGCCCTCAAG 123
DB 112 ACCTTCGCTAGTGAAGGTTCCAAAGGGGAGACCTTATCTGTGTCAACACTGCCCTGAAA 171
QY 124 ACCGTTACCGTCACTTGGACTGTGCTGGTACTACCTGAGCAAGGGTGAGTGGTGAG 183
DB 172 ACCGTTACCGTCACTTGGACTGTGCTGGTACTACCTGAGCAAGGGTGAGTGGTGAG 231
QY 184 GGTATCCGTGACTTCCCTGAAGGAGAACCCCTCGGTGAAGGCTGAGGACATCTTCGTCTGC 243
DB 232 GGTATCCGTGACTTCCCTGAAGGAGAAAACCCCTCGGTGAAGGCTGAGGACATCTTCGTCTGC 291
QY 244 ACCAAGGTGTGGAACCACTCCACCGTTATGAGGACGTTCCTGGTCCATTGACGACTCC 303
DB 292 ACCAAGGTGTGGAACCACTCCACCGTTATGAGGACGTTCCTGGTCCATTGACGACTCC 351
QY 304 CTGAAGCGTCTTGACCTGACTGTGATATGTTCTGTTCACTGGCCCATTTGCTGCC 363
DB 352 CTGAAGCGTCTTGACCTGACTGTGATATGTTCTGTTCACTGGCCCATTTGCTGCC 411
QY 364 GAGAAGATGCGCAGGGTGAGCCCAAGATTGGCCCTGACGCCAAATACGTCATTCTCAAG 423
DB 412 GAAAAAATGCGCAGGGTGAGCCCAAAATTTGGCCCTGACGGCAAAATACNTCTCTCAAG 471
QY 424 GACCTGACGAGACCCCGACCCACATGCGCGCTATGAGAGATTTATGAGATCGC 483
DB 472 GACCTGACCGA-AAANCCCNANCCACCTGGCGCTATGGAAGAAAATTTTNGANGATCCC 530

QY 484 AAGCCAGGTCCATTGGTGTCTCCAACTGGACCATTCGCCACCTTGAGAGATGTCCAAAG 543
DB 531 AAGCCAGGTCCATTGGTGTCTCCAACTGGACCATTCGCCACCTTGAGAGATGTCCAAAG 590
QY 544 TTGCGCAAGGTCAATCCCTCAGCCCAACAGATCGAGATTCACCCCTTCTGCGCCCAACGAG 603
DB 591 TTGCGCAAGGTCAATCCCTCAGCCCAACAGATCGAGATTCACCCCTTCTGCGCCCAACGAG 650
QY 604 GAGCTGCTGCTAGTACTGCTTCTCCAGAACATTTATGCCCGTGGCCTA 650
DB 651 GAGCTGCTGCTAGTACTGCTTCTCCAGAACATTTATGCCCGTGGCCTA 697

RESULT 6

US-10-004-115A-22
Sequence 22, Application US/10004115A
Publication No. US20030134402A1
GENERAL INFORMATION:
APPLICANT: ASAKO, HIROYUKI
APPLICANT: MATSUMURA, KENJI
APPLICANT: SHIMIZU, MASATOSHI
APPLICANT: ITO, NOBUYA
APPLICANT: WAKITA, RYUHEI
TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
FILE REFERENCE: 7372-72249
CURRENT APPLICATION NUMBER: US/10/004,115A
CURRENT FILING DATE: 2002-10-23
PRIOR APPLICATION NUMBER: JP 2000-372704
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: JP 2001-006144
PRIOR FILING DATE: 2001-01-15
PRIOR APPLICATION NUMBER: JP 2001-026594
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: JP 2001-175175
PRIOR FILING DATE: 2001-06-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 22
LENGTH: 417
TYPE: DNA
ORGANISM: Escherichia coli
US-10-004-115A-22

Query Match 42.6%; Score 417; DB 15; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.9e-123;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 CCAAGGTGTGGAACCACTCCACCGTTATGAGGACGTCTCTGGTCCATTGACGACTCCC 304
DB 1 CCAAGGTGTGGAACCACTCCACCGTTATGAGGACGTCTCTGGTCCATTGACGACTCCC 60
QY 305 TGAAGCGTCTTGGACTTGACTACGTGTATATGTTCTCGTTCACCTGGCCCATTTGCTGCCG 364
DB 61 TGAAGCGTCTTGGACTTGACTACGTGTATATGTTCTCGTTCACCTGGCCCATTTGCTGCCG 120
QY 365 AGAAGATGCGCAGGTCAGCCCAAGATTCGCTTGAAGCAAAATACGTCATTCTCAAGG 424
DB 121 AGAAGATGCGCAGGTCAGCCCAAGATTTGGCCCTGAGCGCAAAATACGTCATTCTCAAGG 180
QY 425 ACCTGACCGAGAACCCCGAGCCACATGGCGGCTATGAGAGAAATTTATGAGGATCGCA 484
DB 181 ACCTGACCGAGAACCCCGAGCCACATGGCGGCTATGAGAGAAATTTATGAGGATCGCA 240
QY 485 AGCCAGGTCCTATGGTGTCTCCAACTGGAACATTCGCGACCTTGAGAGATGTCCAAAGT 544
DB 241 AGCCAGGTCCTATGGTGTCTCCAACTGGAACATTCGCGACCTTGAGAGATGTCCAAAGT 300
QY 545 TCGCCCAAGGTCTATGCTCTACGCCAACACAGATCGAGATTCACCCCTTCTGCGCCCAACGAGG 604
DB 301 TCGCCCAAGGTCTATGCTCTACGCCAACACAGATCGAGATTCACCCCTTCTGCGCCCAACGAGG 360

241	TGCAACAAGGTGTGGAACCACTCCACGGTTATGAGGACGTCTCTGCTCCATTGACGAC	300
Db		
301	TCCCTGAAGCGCTTTGGACTTGACTACGTTGATATGTTCTCTGTTCACTGGCCATTGCT	360
Qy		
301	TCCCTGAAGCGCTTTGGACTTGACTACGTTGATATGTTCTCTGTTCACTGGCCATTGCT	360
Db		
361	GCGGAGAGAATGCCACGGGTGACCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC	420
Qy		
361	GCGGAGAGAATGCCACGGGTGACCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC	420
Db		
421	AAGGACCTGACCCGAGAACCCCGAGCCCAACATGCGCGCTATGAGAGAGATTTATGAGGAT	480
Qy		
421	AAGGACCTGACCCGAGAACCCCGAGCCCAACATGCGCGCTATGAGAGAGATTTATGAGGAT	480
Db		
481	CGCAAGCCAGGTCCATTGGTGTCTCCAACTGGACCAATTGCCGACCTTGAGAAGAGATGTC	540
Qy		
481	CGCAAGCCAGGTCCATTGGTGTCTCCAACTGGACCAATTGCCGACCTTGAGAAGAGATGTC	540
Db		
541	AAGTTTCGCCAAGGTCAATGCCCTCAGCGCAACCAAGATCGAGATTCACCCCTTCCTGCCCAAC	600
Qy		
541	AAGTTTCGCCAAGGTCAATGCCCTCAGCGCAACCAAGATCGAGATTCACCCCTTCCTGCCCAAC	600
Db		
601	GAGGAGCTGGTGACGACTCTCTTCTCCAGAACTATATGCCGTGGCCTACTCTCTCTG	660
Qy		
601	GAGGAGCTGGTGACGACTCTCTTCTCCAGAACTATATGCCGTGGCCTACTCTCTCTG	660
Db		
661	GGCTCGCAGAACCCAGGTTCCCAACACCGGTGAGCGGTACGCGAGAACAGACTCTTGAAC	720
Qy		
661	GGCTCGCAGAACCCAGGTTCCCAACACCGGTGAGCGGTACGCGAGAACAGACTCTTGAAC	720
Db		
721	GAGATCGCGGAGAGGGGGGCAACACCTTGTCTCAGGTTCTTATTCGCTGGGGTCTGGCG	780
Qy		
721	GAGATCGCGGAGAGGGGGGCAACACCTTGTCTCAGGTTCTTATTCGCTGGGGTCTGGCG	780
Db		
781	CGTGGCTAGTCTGTTCTCCCCAAGAGCTCCAAACCCCAAGCGCATTTGAGTCCAACTTCAAG	840
Qy		
781	CGTGGCTAGTCTGTTCTCCCCAAGAGCTCCAAACCCCAAGCGCATTTGAGTCCAACTTCAAG	840
Db		
841	AGCATTGAGCTCTCGATGCGCATCTTTGAGCCATCAATGCCGTTGCCAAGGGTCTGTCAC	900
Qy		
841	AGCATTGAGCTCTCGATGCGCATCTTTGAGCCATCAATGCCGTTGCCAAGGGTCTGTCAC	900
Db		
901	TTCCGTTTCGTCAACATCAAGCATACTTTTGGATATGATGCTCTGGCCCGAGGAGACCGGC	960
Qy		
901	TTCCGTTTCGTCAACATCAAGCATACTTTTGGATATGATGCTCTGGCCCGAGGAGACCGGC	960
Db		
961	AAGAACCTGTTCTGGGTGA	978
Qy		
961	AAGAACCTGTTCTGGGTGA	978
Db		

RESULT 3
US-10-608-533-2

Qy	721	GAGATCGCGCAGAGAGGGCGGCAACACCGTTGCTCAGGTTCTTATTTGCTGGGGTCTGGCG	780
Db	721	GAGATCGCGCAGAGAGGGCGGCAACACCGTTGCTCAGGTTCTTATTTGCTGGGGTCTGGCG	780
Qy	781	CGTGGCTAGCTGTTCTCCCCAAGAGGTCCTCAACCCCAAGCGCATTTGAGTCCAACTTCAAG	840
Db	781	CGTGGCTAGCTGTTCTCCCCAAGAGGTCCTCAACCCCAAGCGCATTTGAGTCCAACTTCAAG	840
Qy	841	AGCATTTAGGCTCTCGATGCGCATTTTGAGGCCATCAATGCCGTTGCCAAGGGTCTGTAC	900
Db	841	AGCATTTAGGCTCTCGATGCGCATTTTGAGGCCATCAATGCCGTTGCCAAGGGTCTGTAC	900
Qy	901	TTCCGTTTCGTCAACATGAAGGATACTTTGGGATATGATGCTGGCCCGAGGAGACCGCC	960
Db	901	TTCCGTTTCGTCAACATGAAGGATACTTTGGGATATGATGCTGGCCCGAGGAGACCGCC	960
Qy	961	AAGAACCCTGTCTGGGTGA	978
Db	961	AAGAACCCTGTCTGGGTGA	978

Query Match

RESULT 4
US-10-004-115A-28
; Sequence 28, Application US/10004115A
; Publication No. US2003013402A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: MATSUMURA, KENJI
; APPLICANT: SHIMIZU, MASATOSHI
; APPLICANT: ITO, NOBUYA
; APPLICANT: WAKITA, RYUHEI
; TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
; TITLE OF INVENTION: 4-HALO-3-HYDROXYBUTANOATE
; FILE REFERENCE: 7372-72249
; CURRENT APPLICATION NUMBER: US/10/004, 115A
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: JP 2000-372704
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: JP 2001-006144
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: JP 2001-025594
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: JP 2001-175175


```
QY 1 ATGTCTAACGGAAGACTTTTACATTGAGCAACGGCGTCAAGATTCTCGCGTGGCTTT 60
Db 1 ATGTCTAACGGAAGACTTTTACATTGAGCAACGGCGTCAAGATTCTCGCGTGGCTTT 60
QY 61 GGTACCTCGCTAGTGAAGTTTCAAGGGCGAGACTTACTCTGTGCACCACTGGCCCTG 120
Db 61 GGTACCTCGCTAGTGAAGTTTCAAGGGCGAGACTTACTCTGTGCACCACTGGCCCTG 120
QY 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 GAGGGTATCGGTGACTTCTGAGGAGAACCCCTCGGTGAGCGTGAAGCACTTCGTC 240
Db 181 GAGGGTATCGGTGACTTCTGAGGAGAACCCCTCGGTGAGCGTGAAGCACTTCGTC 240
QY 241 TGCACCAAGGTGTGGAACCACTTCCACCGTTATGAGGACGTCCTCTGCTGCTGCTGCT 300
Db 241 TGCACCAAGGTGTGGAACCACTTCCACCGTTATGAGGACGTCCTCTGCTGCTGCTGCT 300
QY 301 TCCCTGAAGCGTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 TCCCTGAAGCGTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GCGGAGAAAGTATGCGGAGTGAAGTATGCGGCTGAGGACGTCCTCTGCTGCTGCTGCT 420
Db 361 GCGGAGAAAGTATGCGGAGTGAAGTATGCGGCTGAGGACGTCCTCTGCTGCTGCTGCT 420
QY 421 AAGACCTGACCGAGAACCCCGAGCCCAATGAGCGGCTATGAGGAAAGATTATGAGGAT 480
Db 421 AAGACCTGACCGAGAACCCCGAGCCCAATGAGCGGCTATGAGGAAAGATTATGAGGAT 480
QY 481 CGCAAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 CGCAAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 AAGTTCGCAAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 AAGTTCGCAAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GCGTTCGCAAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GCGTTCGCAAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 GAGATCGCGGAGAGGGGGGCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 GAGATCGCGGAGAGGGGGGCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 CGTGGCTACGTCGCTTCTCCCAAGAGCTTCCAAACCCCAAGCGCATTTGAGTCCAACTTCAAG 840
Db 781 CGTGGCTACGTCGCTTCTCCCAAGAGCTTCCAAACCCCAAGCGCATTTGAGTCCAACTTCAAG 840
QY 841 AGCATTTGAGTCTCGGATGCGGATTTTGAAGGCAATCAATGCGGTTGCGCAAGGGTCTGAC 900
Db 841 AGCATTTGAGTCTCGGATGCGGATTTTGAAGGCAATCAATGCGGTTGCGCAAGGGTCTGAC 900
QY 901 TTCCGTTTCGTCATCAATGAGGATTTTGAAGGCAATCAATGCGGTTGCGCAAGGGTCTGAC 960
Db 901 TTCCGTTTCGTCATCAATGAGGATTTTGAAGGCAATCAATGCGGTTGCGCAAGGGTCTGAC 960
QY 961 AAGAACCTGTCTGGGTGA 978
Db 961 AAGAACCTGTCTGGGTGA 978
```

RESULT 2

US-10-327-108-4

; Sequence 4, Application US/10327108

```
; Publication No. US20030186400A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: WAKITA, RYUHEI
; APPLICANT: ITOH, NOBUVA
; TITLE OF INVENTION: METHOD FOR PRODUCING OPTICALLY ACTIVE
; TITLE OF INVENTION: 2-HYDROXYCYCLOALKANECARBOXYLIC ACID ESTER
; FILE REFERENCE: Q73517
; CURRENT APPLICATION NUMBER: US/10/327,108
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: JP 2001-395884
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: JP 2001-395885
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: JP 2002-107648
; PRIOR FILING DATE: 2002-04-10
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Penicillium citrinum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(978)
US-10-327-108-4
```

```
Query Match 100.0%; Score 978; DB 15; Length 978;
Best Local Similarity 100.0%; Pred. No. 1e-303;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTCTAACGGAAGACTTTTACATTGAGCAACGGCGTCAAGATTCTCGCGTGGCTTT 60
Db 1 ATGTCTAACGGAAGACTTTTACATTGAGCAACGGCGTCAAGATTCTCGCGTGGCTTT 60
QY 61 GGTACCTCGCTAGTGAAGTTTCAAGGGCGAGACTTACTCTGTGCACCACTGGCCCTG 120
Db 61 GGTACCTCGCTAGTGAAGTTTCAAGGGCGAGACTTACTCTGTGCACCACTGGCCCTG 120
QY 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 GAGGGTATCGGTGACTTCTGAGGAGAACCCCTCGGTGAGCGTGAAGCACTTCGTC 240
Db 181 GAGGGTATCGGTGACTTCTGAGGAGAACCCCTCGGTGAGCGTGAAGCACTTCGTC 240
QY 241 TGCACCAAGGTGTGGAACCACTTCCACCGTTATGAGGACGTCCTCTGCTGCTGCTGCT 300
Db 241 TGCACCAAGGTGTGGAACCACTTCCACCGTTATGAGGACGTCCTCTGCTGCTGCTGCT 300
QY 301 TCCCTGAAGCGTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 TCCCTGAAGCGTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GCGGAGAAAGTATGCGGAGTGAAGTATGCGGCTGAGGACGTCCTCTGCTGCTGCTGCT 420
Db 361 GCGGAGAAAGTATGCGGAGTGAAGTATGCGGCTGAGGACGTCCTCTGCTGCTGCTGCT 420
QY 421 AAGACCTGACCGAGAACCCCGAGCCCAATGAGCGGCTATGAGGAAAGATTATGAGGAT 480
Db 421 AAGACCTGACCGAGAACCCCGAGCCCAATGAGCGGCTATGAGGAAAGATTATGAGGAT 480
QY 481 CGCAAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 CGCAAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 AAGTTCGCAAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 AAGTTCGCAAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2004, 11:34:38 ; Search time 495 Seconds
(without alignments)
8987.526 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978

Sequence: 1 atgtctaacgaaagacttt.....ccaagaacctgtctgcgtga 978

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2860401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters: 5920802

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	978	100.0	978	15	US-10-004-115A-2
2	978	100.0	978	15	Sequence 2, Appli
3	978	100.0	978	17	Sequence 4, Appli
4	978	100.0	996	15	Sequence 2, Appli
5	588.6	60.2	697	15	US-10-004-115A-28
6	417.42	42.6	417	15	US-10-004-115A-15
7	412	42.1	743	15	US-10-004-115A-22
8	408.2	41.7	765	16	US-10-004-115A-19
9	265	27.1	331	15	US-10-369-493-27083
10	199.6	20.4	972	13	US-10-004-115A-18
11	144.2	14.7	1121	14	US-09-734-237B-31
12	142.6	14.6	1077	14	US-10-040-416-5
13	130.8	13.4	1289	13	US-10-040-416-3
14	129.4	13.2	1289	13	US-10-425-114-17144
					Sequence 3, Appli
					Sequence 17144, A
					Sequence 25069, A

15 129 13.2 1185 13 US-10-425-114-30950
16 129 13.2 1291 13 US-10-425-114-20425
17 119.2 12.2 1335 17 US-10-641-643-1010
18 119.2 12.2 1367 9 US-09-864-864-332
19 119.2 12.2 1367 13 US-10-342-887-641
20 119.2 12.2 1367 13 US-10-172-118-641
21 119.2 12.2 1367 15 US-10-171-581-1
22 113.2 11.6 942 13 US-09-734-237B-43
23 113 11.6 1337 13 US-10-205-331-29
24 113 11.6 1339 12 US-10-152-319A-1492
25 113 11.6 1339 16 US-10-191-803-12
26 111.8 11.4 1313 9 US-09-925-301-47
27 111 11.3 972 15 US-10-321-204-11
28 111 11.3 972 15 US-10-321-204-13
29 111 11.3 972 15 US-10-321-204-16
30 111 11.3 1290 15 US-10-321-204-8
31 111 11.3 1290 15 US-10-198-846-12927
32 110.8 11.3 1171 13 US-10-425-114-5041
33 110.8 11.3 1226 13 US-10-425-114-15152
34 109.4 11.2 1064 15 US-10-321-204-12
35 109.4 11.2 1194 15 US-10-321-204-15
36 109.4 11.2 1219 13 US-10-403-161-85
37 109.4 11.2 1219 15 US-10-321-204-10
38 109.4 11.2 1225 13 US-10-403-161-91
39 109.4 11.2 1596 13 US-10-116-802-150
40 108.6 11.1 1006 13 US-10-114-270-147
41 108.2 11.1 972 13 US-09-734-237B-33
42 106.6 10.9 1119 9 US-09-850-716A-434
43 106.6 10.9 1119 9 US-09-897-778-434
44 106.6 10.9 1119 13 US-10-007-700-434
45 106.6 10.9 1119 15 US-10-117-982-434

ALIGNMENTS

RESULT 1

US-10-004-115A-2
; Sequence 2, Application US/10004115A
; Publication No. US20030134402A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: MATSUMURA, KENJI
; APPLICANT: SHIMIZU, MASATOSHI
; APPLICANT: ITO, NOBUYA
; APPLICANT: WAKITA, RYUHEI
; TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
; FILE REFERENCE: 4-HALO-3-HYDROXYBUTANOATE
; CURRENT APPLICATION NUMBER: US/10/004,115A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: JP 2000-372704
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: JP 2001-006144
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: JP 2001-026594
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: JP 2001-175175
; PRIOR FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Penicillium citrinum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(975)
US-10-004-115A-2

Query Match 100.0%; Score 978; DB 15; Length 978;
Best Local Similarity 100.0%; Pred. No. 1e-303;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
Qy 323 ACTAGTTGATATGTTCTCTGTTCACTGGCCCATTTGCTGCCGAGAGATGGCCAGGGTG 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 380 AGTACCTCGATCTGTACCTTATCCACTGGCCCATGGGCTACAGGAG-----G 427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 383 AGCCCAAGATTGGCCCTGACGGCAAAATACGTCAATTCTAAGGACCTGACCGAGAACCCCG 442
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 428 GATGGATCTGTTCCCAACGACAAAGGATGGCAAGACGCTGTACTCGCCGGTTGATTACG 487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 443 AGCCCAATGGCGGCTATGAGAGAGATTTATGAGGATCGCAAGGCCAGGTCCATTGGTG 502
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 488 TCGACACGTGGAAGCCATGGAGAAGTTGGTGAAGAGGGTCTGGTCAAGTCCATTGGTG 547
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 503 TCTCAACTGGACCATTTGCCACCTTGAGAAGATGTCCAAAGTTGCCCAAGGTCATGCCTC 562
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 548 TTTCCAACTTCAACAGAAGGCAGATCGAGCGCGTGCTTGAGGTGGCCCACTATTCCACCAG 607
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 563 AGGCCAACCCAGATCGAGATTACCCCTTCTGCCCCACGAGGAGCTGGTGCAGTACTGCT 622
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 608 TAACCAATCAGATTGAGTGGCCATCCATATCTGACCCAGAAAGAGCTGATTGACTTCTGCA 667
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 623 TCTCAAGAACATTATGCCCGTGGCCTACTCTCCTCTGGGCTCGCAGAACACAGTTCCCA 682
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 668 AGTCAAGACATTACATCACTGCCTACAGTCCCTTGGATCTCCCAACCCGCATGGGCA 727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 683 CCACCGGTGACGGGTACGCGAGAACAG 711
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 728 AGGCTGGTGATACGGTCACTCTAGAGGAGG 756
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: May 29, 2004, 12:59:21
Job time : 2529 secs

JOURNAL
COMMENT

Unpublished (2002)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 638 Std Error: 0.00
Plate: 13 row: N column: 5
Seq primer: M13rev.

FEATURES
source

Location/Qualifiers
1..638
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:263815"
/db_xref="taxon:112509"
/clone="H113N05"
/tissue_type="endosperm early"
/dev_stage="0-16 hours after imbibition"
/lab_host="X110-Gold"
/clone_lib="H110"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

ORIGIN

Query Match 15.2%; Score 148.8; DB 13; Length 638;
Best Local Similarity 55.1%; Pred. No. 3.6e-22;
Matches 335; Conservative 0; Mismatches 267; Indels 6; Gaps 2;

QY 59 TTGTTACCTTCGCTAGTGAAGGTTCCAAAGGCGAGACCTATACGCTGTCCACTGCC 118
DB 15 TTGTTTCGGACATGGAAGCTAGCTGTGCTGAGTGGCGCTGCTGCTGAGGCC 74
QY 119 TGAAGACCGTTACCGTCACTTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 178
DB 75 TCAAGCTGGCTACCGCCACCTTGTATCTGCGCAAGGCTCTACGGCAACCAAGAGATCG 134
QY 179 GTGAGGGTATCCGTGAC---TTCCTGAGGAGAACCCCTCGTGAACGCTGAGACATCT 235
DB 135 CCCAGGCCCTCAAGAAGGCTTTGCGCGCGAGGTCCCTTAACCTCAAGCGTGAGACGTCT 194
QY 236 TCGTCTGCCCAAGGTGTGGAACCAACCTCCACCGCTTATGAGGACGCTCTCTGCTCCATTG 295
DB 195 TCATCACATCAAGCTGTGGAACCTGCGAGCACCGCCCAAGGACGTGCGCGGCTCTCG 254
QY 296 AGGACTCCTGAAGCGCTTGGACTGACTGACTGATGATGTTCTGCTGCTGCTGCTGCTG 355
DB 255 ACGACTGCTGGCTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
QY 356 TTGCTGCCGCAAGAAATGCGCAGGCTGAGCCCAAGATTGGCCCTGAGCGCAAAATACGTCA 415
DB 315 TCGCTTTTCGAGCGCGCGCGATGTCCACACACCTCTTCCCTCTCGAGGGCGAGAGCG 374
QY 416 TTCTCAAGACCTGACCGAGAACCCCGAGCCCAATGCGCGCTGATGAGAAATTTATG 475
DB 375 TCAAGATCTCGACACGCTCTCTATGCTGACATGGAAGGCCATGACAGAGTGGCCA 434
QY 476 AGGATCGCAAGCCAGGTCCTATGTTGCTCCAACTGGACCATTTGCGACCTTCGAGAAGA 535
DB 435 AGGA---GAAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491
QY 536 TGTCCAAAGTTCGCCAAGGTCATGCTCAGCCCAACCAAGATCGAGATTCACCCCTTCCTGC 595
DB 492 TCATCGAGCGGCAACCGCGCTCACTCCCGCTGCGCAACCAAGATCGAGCGCCACCCCGTCTCC 551

QY

596 CCAACGAGGAGCTGCTGAGTACTGTTCTTCCAAAGCAATTATGCCGTGGCTACTCTC 655
DB 552 TCACGCCAGAGCTCATCCAGTACTCAAGGAGAGAACATCCACATCACCAGTACTCCG 611
QY 656 CTCTGGGC 663
DB 612 CCTTCGGC 619

RESULT 15

AA697426

LOCUS

DEFINITION

AA697426

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..758

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="HL02448"

/sex="male and female"

/dev_stage="adult"

/lab_host="SOLR"

/note="Organ: head-brain & sensory organ; Vector: Bluescript SK; Site 1: EcoRI; Site 2: XhoI; Constructed using Stratagene ZAP-cDNA Synthesis kit. Oligo 87-primed and directionally cloned at EcoRI and XhoI in Bluescript SK(+/-)"

ORIGIN

Query Match 15.2%; Score 148.6; DB 9; Length 758;

Best Local Similarity 54.8%; Pred. No. 4.4e-22;

Matches 345; Conservative 0; Mismatches 269; Indels 15; Gaps 2;

QY 83 CCAAGGGCGAGACCTATCTGCTGTCACCATGSCCTGAAGACCGGTTACGCTACTGG 142

DB 143 CCAAGGGTCAGGTCAACGAGGCTGTCAAAGTTGCCATTGATCCCGATACAGGCATTTG 202

QY 143 ACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202

DB 203 ACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 259

QY 203 AGGAGAACCCCTCGGTGAAGCGGTGAGGACATCTTCTGCTGACCAAGGTGTGGAACACC 262

DB 260 TCAAGAGGGCGGTGTCAAGCGGTGAGGATCTGTTATCATCACCAGCAAACTGTGGAACACTT 319

QY 263 TCACCGCTTATGAGGACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322

DB 320 TCATGCTCCCGGCTCTTGTCAAGTGGCATTTGGAGAACATTTGAGCTCCCTGGAAGCTGA 379

Db ACCCTGGCCGAGCTTGCGCTCGACTACCTTGACCTTACCTCATCCACTGCGCCGTTGCC 448
QY GCCGAGAGATGCGCGAGGTGAGCCCAAGATTGGCCCTGACCGCAATACGTCATTCTC 420
Db TTTGCTCCGCGCGGACCTCTCCCAAG--TCCGAGGACGCGCTCCGAGGTGCGAGCTC 505
QY AAGGACCTGACCGAGAACCCCGAGCCCAATGCGCGCTATGAGAGAGATTATGAGGAT 480
Db AA--CCCAAGATGTCTCATTTGTCAGACCTGGAAGCCATGACCGAGCTGCCCAAGTC- 562
QY CGGAGGCGAGTGCATTTGCTGCTCCCACTGAGGACCATGCGGACCTTGAAGATGTC 540
Db --CAAGTCCCTCCGTCGGTGTCTCCCACTTTACCATTTAGACACCCCGAGCGCGCATC 620
QY AAGTTGCGCAAGGTGATGCTTCAGCCCAACAGATGAGATTTACCCCTTCTGCGCCAAC 600
Db GAGGCCACCGCGTCTGCTCCCGCGTAAACAGATCGAGCGCACACCCGCTTCCCAAC 680
QY GAGGAGCTGTGAGTACTGCTTCTCCAGAACATTAT 638
Db CAGCCCTGATCGACTACTGCGCCAGAACGCGATCAT 718
RESULT 11
LOCUS CB899137 823 bp mRNA linear EST 02-JUL-2003
DEFINITION trico17xh08 T.reesei mycelial culture, Version 3 april Hypocrea
jecorina cDNA clone trico17xh08, mRNA sequence.
ACCESSION CB899137
VERSION CB899137.1 GI:30113795
KEYWORDS EST.
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM Hypocrea jecorina
REFERENCE
AUTHORS Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,
Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,
Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,
Olivares, H.A., Reunissen, P.J., Yao, J., and Ward, M.
TITLE Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus Trichoderma reesei
JOURNAL J. Biol. Chem. 278 (34), 31988-31997 (2003)
MEDLINE 22803314
PUBMED 12788920
COMMENT Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-F1 primer.
FEATURES
Source
1..823
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="trico17xh08"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 3 april"
/note="Vector: pREP3v; Site 1: Not 1/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
ORIGIN
Query Match 16.4%; Score 160.4; DB 14; Length 823;
Best Local Similarity 57.7%; Pred. No. 1.1e-24;
Matches 368; Conservative 0; Mismatches 256; Indels 14; Gaps 4;
QY 1 ATGTCTAACGGAAGACTTTTACATTTGAGCAACGCGTCAAGATTCTCGGTCGGCTTT 60
Db 149 ATGTCTTCCGGAAGACCGTCAACCTTCAACCCGCTACAGATCCCGCATCGGTAC 208

QY 61 GTTACCTTCCTAGTGAAGTTTCCAAAGGCGAGAACTATATCTGCTGTCACTACCTGCCCTG 120
Db GGACCT-----GGAGCCCTCCGGCGAGGTGCGGGCTGGTGTCTTTGAGCCCTC 262
QY 121 AAGACCGTTACCGTACCTTGGACTGTGCTGGTACTACCTGAACGAGGGTGAAGTTGGT 180
Db AAGTTGGCTACCGGCACCTCGACCTGGCCAAAGTCTACGGCAACGAGAAGAGGTTGGT 322
QY 181 GAGGTTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAAGACATCTTCGTC 240
Db GAGGGATCAAGAAGGCTTTGCTGAGGTCCCGCCCTGAAGCGGAGGATATTTCATC 382
QY 241 TCACCAAGTGTGAACCACTCCACCGTTATGAGGACGTCCTCTGTTCCATTGACGAC 300
Db ACTTCAAGCTGTGAACAACCTCCCAAGCCGAGGACGTCGAGCCGCTCTCGACGAC 442
QY 301 TCCTGAAGCGTCTGGACTTGAAGTGTGATATGTTCTCTGTTCACTGGCCCAATTGCT 360
Db ACCCTGGCGAGCTTGGCTTCGACTACCTTGAACCTTACCTTACCTCATCTGCGCCCTTGGC 502
QY 361 GCGGAGAGATGCGCAGGTTGAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC 420
Db TTTGCTCCCGCGCGACCTCTTCCCAAG---TCGAGGACGGCTCCGAGGTGCGAGTCTC 559
QY 421 AAGGACCTGACGAGAACCCCGAGCCACATGCGCGCTATGAGAGAGATTATGAGAT 480
Db AA--CCCAAGATGTGTCCATTGTCCAGACCTGGAAGGCGCATGACGAGTGTGCCAAGTC- 616
QY 481 CGCAAGCGCAGTTCATTGTTGTTCTTCAACTGGACCATTTGCCGACTTGAAGAATGTC 540
Db --CAAGTCCGCTCGCTGGTGTCTTCCCACTTACCATTTAGCAACCCGCGCCGATC 674
QY 541 AGTTGCGCAAGTGTATGCTTCAAGCAACAGATGAGATTCAGCCCTTCTGCGCCCAAC 600
Db GAGGCCACCGCGTCTGTCGCCCGCGTAAACACAGATCGAGCGCACACCCGCTTCCCAAC 734
QY 601 GAGGAGCTGTGAGTACTGCTTCTTCCCAAGAACATTAT 638
Db CAGCCCTGATCGACTACTGCGCCAGAACGCGATCAT 772

AA785751 189 bp mRNA linear EST 31-JUL-1998
hlh04a1.x1 Aspergillus nidulans 24hr asexual developmental and
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
hlh04a1 5', mRNA sequence.
AA785751
AA785751.1 GI:2845919
EST.
ORGANISM Emericella nidulans (anamorph: Aspergillus nidulans)
Emericella nidulans
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
REFERENCE 1 (bases 1 to 189)
AUTHORS Kuper, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
Prade, R., and Roe, B.
TITLE An Aspergillus nidulans EST Database
JOURNAL Unpublished (1998)
COMMENT Other_ESTs: hlh04a1.fl
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: SK
High quality sequence stop: 182.
Location/Qualifiers

FEATURES

```

Db      662 ATCAAGTCCAACTTCGAGGCTTTCATCTGTCTCCCGAGGAGATTAAGTCTCTGGACAGC 721
      883 GTTCCCAAGG 892
      722 CTTCGACGAGG 731

RESULT 9
AA786853
LOCUS   534 bp mRNA linear EST 31-JUL-1998
DEFINITION m6f07a1.f1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone m6f07a1 5', mRNA sequence.
ACCESSION AA786853
VERSION   1
KEYWORDS  Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella.
SOURCE    Emericella nidulans (anamorph: Aspergillus nidulans)
ORGANISM  Emericella nidulans
REFERENCE 1 (bases 1 to 534)
AUTHORS   Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.
TITLE     An Aspergillus nidulans EST Database
JOURNAL   Unpublished (1998)
COMMENT   Other ESTs: m6f07a1.f1
          Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
          Department of Chemistry and Biochemistry
          Advanced Center for Genome Technology, University of Oklahoma
          620 Parrington Oval, Norman, OK 73019, USA
          Tel: 405 325 4912
          Fax: 405 325 7762
          Email: broe@ou.edu
          We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
          Seq primer: SK
          High quality sequence stop: 522.
FEATURES
    source
        1..534
            Location/Qualifiers
                /organism="Emericella nidulans"
                /mol_type="mRNA"
                /strain="FGSC A26"
                /db_xref="taxon:162425"
                /clone="m6f07a1"
                /tissue type="vegetative mycelia, asexual structures"
                /clone lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
                /note="vector: pBluescript SK; Site 1: EcoRI; Site 2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN
    Query Match 16.9%; Score 164.8; DB 9; Length 534;
    Best Local Similarity 77.7%; Pred. No. 9.7e-26;
    Matches 199; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY      1 ATGCTTAACGGAAGACATTTACATTTAGCAACGGCGTCAAGATTCCTGCGTCGCGTTT 60
      279 ATGCTTAACGGAAGACATTTCAAGCTCAGCAACGGCGTCAAGATTCCTGCGTCGCGTTT 338
      61 GGTACCTTCGCTAGTGAAGTTTCAAGGGGAGACCTATCTGCTGCACACTGCCTCG 120
      339 GGTACCTTCGCGACGGAGGGCGCTTCGGTGAGACCTACCGCGCGTCAAGAAGCCCTC 398
      121 AAGACCGGTTACCGTCACTTGGACTGTCCTGTTACTTACTTGAACGAGGTGAGGTTGGT 180
      399 GAGGTGCGATACAGCACTTGCAGCTGCGCTGTTTCTTACAGACGAGGACGAGGTTGGC 458
      181 GAGGTATCCGTGATTCCTCAAGGAGAACCCCTCGGTGAGCGTGAGGACATCTTCGTC 240
      459 GATCGCGTTCGCGACTTCCTCAAGGAGAACCCCTCTGTCAAGCGCGGAGGACATCTTCATC 518

```

```

QY      241 TGCACCAAGGTGTGGA 256
      519 TGCACCAAGGTGTGGA 534

RESULT 10
CF869025
LOCUS   769 bp mRNA linear EST 31-OCT-2003
DEFINITION trico17xh08.b1 T.reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone trico17xh08, mRNA sequence.
CF869025
VERSION   1
KEYWORDS  Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocremycetidae; Hypocreales; Hypocreaceae; Hypocrea.
SOURCE    Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM  Hypocrea jecorina
REFERENCE 1 (bases 1 to 769)
AUTHORS   Diener, S.E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D., Mitchell, T.K., van Solingen, P., Teunissen, P.J.M., Ward, M. and Dean, R.A.
TITLE     Analysis of the protein processing and secretion pathways in a Trichoderma reesei EST dataset
JOURNAL   Unpublished (2003)
COMMENT   Contact: Ralph A. Dean
          Fungal Genomics Laboratory
          North Carolina State University
          Campus Box 7251, Raleigh, NC 27695, USA
          Tel: 919-513-0020
          Fax: 919-513-0024
          Email: ralph.dean@ncsu.edu
          Seq primer: LT-F1 primer.
          Location/Qualifiers
              1..769
                  /organism="Hypocrea jecorina"
                  /mol_type="mRNA"
                  /strain="QM6a"
                  /db_xref="taxon:51453"
                  /clone="trico17xh08"
                  /dev stage="mycelia"
                  /clone lib="T.reesei mycelial culture, Version 6 October 2003"
                  /note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."

ORIGIN
    Query Match 16.4%; Score 160.4; DB 14; Length 769;
    Best Local Similarity 57.7%; Pred. No. 1.1e-24;
    Matches 368; Conservative 0; Mismatches 256; Indels 14; Gaps 4;

QY      1 ATGCTTAACGGAAGACATTTACATTTAGCAACGGCGTCAAGATTCCTGCGTCGCGTTT 60
      95 ATGCTTTCGGAAGGACCGTCAACCTCAACACCGGCTACAGATCCCGCAGATCGGCTAC 154
      61 GGTACCTTCGCTAGTGAAGTTTCAAGGGCGAGACCTATCTGCTGTCAACACTGCCCTG 120
      155 GGCACCT-----GCGAGCGCGTCCCGCGAGGTTCGCGCTGTGTCTTTGAGGCGCTC 208
      121 AAGACCGGTTACCGTCACTTGGACTGTCTCGTGGTACTACTCTGTAACGAGGTGAGGTTGGT 180
      209 AAGGTGGCTACCGGCACCTCGACCTGGCAAGGTCTACGGCAACCCAGAGGAGGTTGGT 268
      181 GAGGTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
      269 GAGGCGATCAAGAAGGCTCTTGCTGAGGTCCCGCGCTTAAGCGCGAGATATTTTCAATC 328
      241 TGCACCAAGGTGTGGAACCAACCTCCACCGTTATATGAGGACGTCCTCTGTGTCATTGACGAC 300
      329 ACCTCCAGCTGTGGAACAACTCCCAAGCCGAGGACGTCGAGCCGCGCTCTCGACGAC 388
      301 TCCTTGAAGCGCTTTGGACTTGACTACGTTGATATGTTCTCTCGTTCATCTGCGCCCATTCGT 360

```



```

265 ACCGGAAGCTTTGGAAACAACACCGCCCTGAGGAGGTGCTGTGCTCTTGAGGAC 324
301 TCCCTGAAGCGCTTGAGACTGACTAGCTGATGATGCTGCTGCTGCTGCTGCTGCT 360
325 ACCCTCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 384
361 GCCGAGAAGATGCCAGGCTGAGCCCAAGATTGGCCCTGACGCGCAATACGTCATCTC 420
385 TTCAAGAACGAAACACAGCTTTTCCCACTGAAGAGGGCGATGACGCAAGACTGCTCTT 444
421 AAGCACTGACCGAGAACCCGAGCCACATGCGGCGCTATGAGGAAGATTTATGAGGAT 480
445 GACCAGGAGGTGACCCCTCTCCAGACCTGTGAGGCGCTCACCAAGCTGCCCAAGGAG-- 501
481 CGCAAGGCGAGTCCATTGGTGTCTCCAACTGGACCAATTGCCAGCTTGAGGAAGATGTC 540
502 ---AAGTTCCTCCATTGGTGTCTCTTAACCTTCAAGAGAGATGCTTGAGCAAAATCATC 558
541 AAGTTCGCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
559 AAGGACACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
601 GAGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
619 CTTGAGCTTGTCAAGTACCTCCAGGAGAAGGATGCTGCTGCTGCTGCTGCTGCTGCT 678
661 GGCTCGCAGAACACAGG 676
679 GGCAACAACAGCTGGG 694

RESULT 8
CNS06Q60 1087 bp DNA linear GSS 05-JUL-2001
LOCUS T7 end of clone AW0AA003D07 of library AW0AA from strain CLIB 89 of
DEFINITION Yarrowia lipolytica, genomic survey sequence.
ACCESSION AL410278
VERSION AL410278.1 GI:12178444
KEYWORDS GSS.
SOURCE Yarrowia lipolytica
ORGANISM Yarrowia lipolytica
REFERENCE 1 (bases 1 to 1087)
AUTHORS Souciest,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Durjon,B., Durrens,P., Lepingle,A., Illorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nicche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
PERS Lett. 487 (1), 3-12 (2000)
1152876
2 (bases 1 to 1087)
AUTHORS Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynerol,C.,
Artiguenave,F., Wincker,P. and Gaillardin,C.
Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
lipolytica
PERS Lett. 487 (1), 95-100 (2000)
1152892
3 (bases 1 to 1087)
AUTHORS Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqlife@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
```

```

Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared, and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
Location/Qualifiers
1..1087
/organism="Yarrowia lipolytica"
/mol_type="genomic DNA"
/strain="CLIB 89"
/db_xref="taxon:4952"
/clone="AW0AA003D07"
/clone_lib="AW0AA"
/note="end : T7"
<23..>763
misc_feature
/note="similar to Saccharomyces cerevisiae ORF YDR368w [
IPK1 ; strong similarity to members of the aldo/keto
reductase family ]"
/evidence=not_experimental
<23..>733
misc_feature
/note="similar to Saccharomyces cerevisiae ORF YOR120w [
GCY1 ; galactose-induced protein of aldo/keto reductase
family ]"
/evidence=not_experimental
ORIGIN
Query Match 17.0%; Score 166; DB 29; Length 1087;
Best Local Similarity 56.6%; Pred. No. 7.7e-26;
Matches 379; Conservative 0; Mismatches 270; Indels 21; Gaps 3;
QY 223 CGTGAGGACATCTTCGTGTGCACCAAGGTGTGGAAACCACTCCACCCTGTTAGGACCTC 282
DB 83 CGAGAGGAGATCTTCCTGACCACCAAAACTGGAAACAACGCCCGAACCAGAGAGGGCTC 142
QY 283 CTCGTGTCATTGACGACTCCCTGTAAGCGCTTTGAGCTTGACTAGCTTGATGTTCTCCTC 342
DB 143 AAGAAGCAGCTCGACCAAGTCCCTCAAGGAGCTGCAGACCGACTAGCTTGATGCTGCTC 202
QY 343 GTTCACCTGCCCATTTGTCGCGAGAAAGTATGCGGAGGTGAGCCCAAGATTGGCCCTGAC 402
DB 203 ATCCACTGCGCCGTTGTTTTC---AAGACCGCGGAGGAGCTCTTCCCCACAAACCCGAC 259
QY 403 GGCATTAAGTCTCATCTCAAGGACCTGACCGAGACCCGAGCCACATGGCGGCTATG 462
DB 260 GGCATGTCGCCCTTGGCGAGTCTCCTGTGAGG-----GACTTGGAGGCTTTC 310
QY 463 GAGAAGATTATGAGGATGCGCAAGCCAGGTTCATTGTGTCTCCAACTGGACCAATTGCC 522
DB 311 GAGGAGCTCGTGAAGAGGGCAAGGTCAAGTCCATCGGTGCTCCAACTTCAACCGAGGAG 370
QY 523 GACCTTGAGAGATGTCCTCAAGTTCGCAAGGTTCATGCTCAGCCACCAACAGATCGAGAT 582
DB 371 CGAATTGAGAGCTCTCTCAAGACCCCAAGATCCCCCTGCGCTCAACAGATTGAGTAC 430
QY 583 CACCCCTTCTGCCCAAGCAGGAGCTGGTGAGTACTGCTTCTCCAGAACATTATGCC 642
DB 431 CACCCCTACTCTGACGACGCGGCTCTGCGCAAGTACCTTGGAGCAAGAACATTGCTCTC 490
QY 643 GTGGCTTACTCTCTCTGGGCTCGGAGAACACAGGTTCACCAACCGGTCAGGCGGTGAG 702
DB 491 GAGGCTTACTCTCTCTCGGTAAACAACACTTACGGCTTTCCTCC-----CGAGCTATT 541
QY 703 GAGAACAGACTCTGAACAGAGATCGCCGAGAGGGCGCAACACCCCTTGTCTCAGGTCTT 762
DB 542 GATGACGAGGAGTATCGATCTGCGCAAGAGAGAGGCGAGGATGCTGCTGCCCTCAT 601
QY 763 ATTGCTGGGGTCTGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
DB 602 GTCAACTGATCAGATCCAAAGAACCATGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 823 ATTGAGTCCAACTTCAAGAGCATTTGAGCTTCTCCGATCGCGACTTTTGAAGCCATCAATGC 882
```

Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

Contact Dr. Marian Beremand regarding clone availability Included
is the best homolog from a blastx search of Genbank nr 04-09-01
526 28-53 gi|3724334|dbj|BAA33 (AB014493) reductase [Gibberella
zeae]

Seq primer: T3
High quality sequence stop: 130.

FEATURES

source
1..390
Location/Qualifiers
/organism="Fusarium sporotrichioides"
/mol_type="mRNA"
/strain="Tri 10"
/db_xref="taxon:5514"
/clone="ndolofs"
/clone_lib="Fusarium sporotrichioides Tri 10 overexpressed
cDNA library"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5' end of cDNA cloned into EcoRI site of
pBluescript; 3' end of cDNA cloned into XhoI site of
pBluescript"

ORIGIN

Query Match 19.3%; Score 188.6; DB 12; Length 390;
Best Local Similarity 73.1%; Pred. No. 4.7e-31;
Matches 242; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 2 TGTCTAAGGAAAGACTTTCACATTGAGCAACGGCTCAAGATTCCTGGGCTGGCTTG 61
Db 56 TTTCTAACAAGACTGTCTTACCTCAACACGGTGTCAAGATGCCGGTGTGGGCTTG 115
QY 62 GTACCTTCGCTAGTGAAGGTTCCAAAGGGCGAGACTATCTGTGTCCACACTGCCCTGA 121
Db 116 GCACCTTGCTAACGAGGGCGCCAGGGCGACCTACAGGCTGTCTATTCGCCCTCA 175
QY 122 AGACCGGTACCGTCACTTGACATGTGCTGGTACTACCTGAACGAGGCTGAGGTGGTG 181
Db 176 AGACCGGTACCGACACTGCTGCTGGTGTCTACCTGAACGAGGGGATGTTGGCG 235
QY 182 AGGCTATCGTGAAGTCTCTGAAGAGAGAACCCCTCGGTGAAGCGTGAAGCATCTTGTCT 241
Db 236 AGGCGTGGTGAATTTCTGCCGAGAGAACGGCTCAAGCGGAGGATCTTTTATCT 295
QY 242 GCACCAAGGTGTGAACCACTCCACCGTTATGAGGAGCTCTCTGTCTCAATTCAGCACT 301
Db 296 GCACCAAGGTCTGAACCACTCCACGAGCTGAGGAGTCAAGTGTCTATTCGAGAACT 355
QY 302 CCTGTAGCGTCTTGGACTTGACTACCTTGA 332
Db 356 CTCCTCAAGAACTTTGGCGCTGACTACATTGA 386

RESULT 7

BI948347 1061 bp mRNA linear EST 19-OCT-2001
LOCUS HVSM1009C04f Hordeum vulgare spike EST library HVCNDA0012
DEFINITION (Fusarium infected) Hordeum vulgare subsp. vulgare cDNA clone
BI948347
VERSION HVSM1009C04f, mRNA sequence.
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 1061)
Wing, R., Muehlbauer, G.J., Close, T.J., Kleinhofs, A., Wise, R.,
Heinen, S., Begum, D., Yu, Y., Henry, D., Palmer, M.,
Rambo, J., Simmons, J., Fenton, R.D., Malatrasi, M., Choi, B.W.,
Cates, R. and Main, D.
Development of a genetically and physically anchored EST resource

TITLE

JOURNAL COMMENT

for barley genomics: Fusarium infected Morex spike cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 541
Seq primer: AATTAACCTCTACTAAAGG
High quality sequence stop: 785.

FEATURES

source
1..1061
Location/Qualifiers
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSME1009C04f"
/tissue_type="Spike"
/lab_host="TJCL1"
/clone_lib="Hordeum vulgare spike EST library HVCNDA0012
(Fusarium infected)"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Plants were grown at the University of Minnesota in
the GJ Muehlbauer lab; spikes were harvested and snap
frozen at 0, 1, 2, 3, 4, 5, 6, and 8 days after Fusarium
graminearum inoculation (Heinen). In the TJ Close lab at
the University of California, Riverside, total RNA was
prepared from each sample pool, equal quantities of all
eight RNA pools were combined, poly(A) RNA was purified
from the mixture, one primary unamplified cDNA library was
made, and 1 million pfu were in vivo excised to give
pBluescript SK(-) cDNA phagemids (Choi, Fenton,
Malatrasi). Phagemids were plated and picked at the
Clemson University Genomics Institute (CUGI) (Begum,
Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
Oates, Rambo, Main). The sequence has been trimmed to
remove vector sequence and contains a minimum of 100 bases
of phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html")

ORIGIN

Query Match 18.0%; Score 176; DB 12; Length 1061;
Best Local Similarity 56.1%; Pred. No. 4.8e-28;
Matches 379; Conservative 0; Mismatches 285; Indels 12; Gaps 2;
QY 1 ATGCTTAAGGAAGACTTTCACATTGAGCAACGGCTCAAGATTCCTGGGCTGGCTTT 60
Db 31 ATGCTTTCGGTCGAACTGTCACTCTCAACTCGGCTGGAAGATCCCCCAGATGGGTAC 90
QY 61 GGTACTTTCGCTAGTGAAGGTTCCAAAGGGCGAGACTATCTGTGTCCACACTGCCCTG 120
Db 91 GGCACAT-----GGCAAGCTGCTCCCGTGGGTGGCAATGGTGTCTACGAGGCCCTT 144
QY 121 AGACCGGTACCGTCACTTGAGCTGTCCCTGGTACTACTCTGAACGAGGCTGAGGTGGT 180
Db 145 AAGGCTGTCTACCGTCACTTGTCTCCCAAGATCTTACCAGAACCCAGCGTGGGTGGT 204
QY 181 GAGGCTATCCGTGACTTCTCTGAAGAGAAACCCCTCGGTGAAGCGTGAAGCATCTTCGTC 240
Db 205 GAGGCTATCAAGAGGCCCTCAGCGAGTCCCTGGCTTAAGCGTGAGGACATCTTCATC 264
QY 241 TGCACCAAGGTGTGAACACCTCCACCGTTATGAGGAGCTCTCTGTGTCATTGACGAC 300


```

Qy      313  CTTGGACTTGACTAGTTGATATGTTCTCTGTTCACTGGCCCAATGCTGCCGAGAAGAAT 372
Db      463  CTAAGGTCGATTACATTGACCTGTTCTCTGTTCACTGGCCCAATGCTGCCGAGAAGAAT 522
Qy      373  GGCAGAGGTGAGCCCRAGATTGGCCCTGAGCGGCAATACGTCATTCTCAAGGACCTTGACC 432
Db      523  AGGCACAGGAGCGTCAAGCTGGGCCCCGATGGCAAGTATGTCATCAACCAAGCCCTCGACG 582
Qy      433  GAGAACCCCGAGCCACATCGCGGCTATGGAGAAGATTTATGAGGATCGCAAGGCCAGG 492
Db      583  GARAACCCAGAGCCAAATCGCGAGCCATGGAAGAGCTTGTGAAAGCGSCCTCGTCAAG 642
Qy      493  TCCATTGGTGTCTCAACTGACCAATTCGCGACCTTGAGAAGATGTCCAAAGTTGCGCAAG 552
Db      643  GCAATTGGAGTATCCAACTGACGATNCGGGGTGAAGAAGCTCTCTTCAGATCGCCCAAG 702
Qy      553  GTCATGCTCAGCGCAACCAAGATCGAGATCA 584
Db      703  ATCAAGCGCGCAGTGAACCAAGATNGAATCA 734

RESULT 3
CB897574
LOCUS
DEFINITION
tricol1xm03 T.reesei mycelial culture, Version 3 april Hypocrea
jecorina cDNA clone tric011xm03, mRNA sequence.
ACCESSION
CB897574
VERSION
CB897574.1 GI:30112232
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
1 (bases 1 to 792)
Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J.,
Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C.,
Olivares,H.A., Teunissen,P.J., Yao,J. and Ward,M.
Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
JOURNAL
MEDLINE
PUBMED
12788920
COMMENT
Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-F1 primer.
Location/Qualifiers
1..792
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/db_xref="taxon:51453"
/clone="tric011xm03"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 3 april"
/note=vector: pREP3y; Site 1: Not 1/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

FEATURES
source
Query Match 27.1%; Score 265; DB 14; Length 792;
Best Local Similarity 66.3%; Pred. No. 9.9e-48;
Matches 379; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

Qy      13  AAGACTTTCAGTGAACCGGTGCAAGATTCCTGGCTCGGCTTGCTACCTTCGCT 72
Db      221  AAGACGTACACTCTGAACCGGTGCAAGATTCCTGGCTCGGCTTGCTACCTTCGCT 280

```

```

Qy      73  AGTGAGGTTCCAAAGCGGAGACCTATATCTGCTCTCACTGCTCCCTGAAGACCGGTTAC 132
Db      281  AATGAGGTTCCAAAGCGGAGACATACGACATGTTTCAAAAGGCACTGGAGCTGGATAC 340
Qy      133  CGTCACTTGGACTGTGCTGCTTACCTGAACAGAGGTGAGTTGGTGGAGGTTATCCGT 192
Db      341  CGCACCTTGATTGGCGCTGGTTTACCAACAGAGATGAGTTGGTGGAGCGGTACGC 400
Qy      193  GACTTCTTGAAGGAGAACCCCTCGGTGAAGCGTGAAGCATCTTCTGCTGACCAAGGTG 252
Db      401  GATTTCTCGCCCGCGACCGACGCTGAACCGCGAGGATCTCTTCATTGACCAAGTT 460
Qy      253  TGGAACCACTCCACCGTTATGAGGACCTCTCTGCTCCATTGACGACTCCTCTGAAGCGT 312
Db      461  TGGAACCACTCTGATGAGCGAGACGTCAGTGGAGCGCCAGAACTCTGTCGCAAAAC 520
Qy      313  CTTGAGCTTGACTACGTTGATATGTTCTGTTCACTGGCCCAATGCTGCCGAGAAGAT 372
Db      521  CTCAAGGTCGATTACATTGACCTGTTCTGCTGCTCACTGGCCAATCGCGCGCGAGAAG 580
Qy      373  GGCAGAGGTGAGCCCAAGATTGGCCCTGACGCGCAATACGTCATTCTCAAGCACTGACC 432
Db      581  AGCGACAGAGCGTCAAGCTGGGCCCCGATGCGAAGTATGTCATCAACCAAGCCCTGACG 640
Qy      433  GAGAACCCCGAGCCACATGCGCGCTATGGAGAAGATTTATGAGGATCGCAAGGCCAGG 492
Db      641  GAAACCCAGAGCCAAATGCGAGCCATGGAAGAGCTTGTGAAAGCGGCTCGTCAAG 700
Qy      493  TCCATTGGTGTCTCAACTGAGACCAATTCGCCACCTTGAGAGATGTCCAAGTTGCCCAAG 552
Db      701  GCAATTGGAGTATCCAACTGAGACGATNCGGGGTGAAGAAGCTCCTTCAGATCGCCCAAG 760
Qy      553  GTCATGCTCAGCGCAACCAAGATCGAGATCA 584
Db      761  ATCAAGCGCGCAGTGAACCAAGATNGAATCA 792

CF866105 704 bp mRNA linear EST 31-OCT-2003
tric005xa21.bl T.reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone tric005xa21, mRNA sequence.
CF866105
CF866105.1 GI:38120731
ACCESSION
VERSION
KEYWORDS
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
1 (bases 1 to 704)
Diener,S.E., Dankmeyer,L., Dunn-Coleman,N., Houfek,T.D.,
Mitchell,T.K., van Solingen,P., Teunissen,P.J.M., Ward,M. and
Dean,R.A.
Analysis of the protein processing and secretion pathways in a
Trichoderma reesei EST dataset
Unpublished (2003)
JOURNAL
COMMENT
Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu
Seq primer: LT-F1 primer.
Location/Qualifiers
1..704
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/db_xref="taxon:51453"
/clone="tric005xa21"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
2003"

FEATURES
source
Query Match 27.1%; Score 265; DB 14; Length 792;
Best Local Similarity 66.3%; Pred. No. 9.9e-48;
Matches 379; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```

Email: r.bruskiewich@giar.org
International Rice Information System (IRIS);
http://www.iris.irri.org; DO200799
Assignment of putative function to the sequence by S. Rudd of the
Munich Information Center for Protein Sequences
(http://mips.gsf.de)
Plate: 03 row: B column: 08.
Location/Qualifiers
1. 886
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR64"
/db_xref="taxon:39946"
/clone="C0000800"
/tissue_type="panicles"
/dev_stage="Flowering"
/clone_lib="IRRI Drought Stress Panicle Library"
/note="Vector: pbluescript II SK+; Water stress was
applied by not watering for 4 consecutive days. Panicles
were collected from control (well watered) and stressed
plants at 2 days before heading, at heading, 50% flowering
and 4 days after 50% flowering."

FEATURES
source

ORIGIN
Query Match 36.1%; Score 353.2; DB 14; Length 886;
Best Local Similarity 68.2%; Pred. No. 3.8e-67;
Matches 527; Conservative 0; Mismatches 242; Indels 4; Gaps 3;
QY 5 CTAACGGAAGACTTTCACATGTAGCAACGCGTCAAGATTCCTGGCGTTGGTA 64
DB 114 CAACACGAGACATACAAAGCTCAACATGCGTTGAGATTCCTGGCGTTAGGATTCGAA 173
QY 65 CTTTCGCTAGTAGAGTTTCAAGGGCGAGACCTATCTGCTGACCACTGCCCTGAAGA 124
DB 174 CATTGCAACAGAGGCTTCAAGGGCGAGACCTACGCTGCGCTCACATGCTCTGAAGA 233
QY 125 CCGGTACCGTCACTTGGACTGTCCCTGGTACTACCTGAACGAGGTTGAGTTGGTAGG 184
DB 234 CTGGCTACCGACACTCGACTGCGATGTTTCTACCAAGACGAGGCGAGTGGTGGAGG 293
QY 185 GTATCCGTGACTTCTGAGAGAGACCCCTCGTGAAGCGTGAAGACATCTTCCTGCGCA 244
DB 294 CCGTCCGGGACTTCTTCGCGCCAGAACTCCCTCCGTCATCTCGCAGGACATCTTATCTGCA 353
QY 245 CCAAGGTGTGAACACACTCCACCGCTTATGAGGACGCTCTCTGTCTCAATGAGACTCCC 304
DB 354 CAAGGTCTGAACCACTCTCACGAGCCGAGGAGTAGATGTCTGTTCAACAGCTCGC 413
QY 305 TGAAGCGTCTTGGACTTGACTAGCTGATGATGTTCTCTGTTCACTGGCCCATTTGCTCCG 364
DB 414 TCGAAGATTTCAGATGGACTACATCGATCTCTTTCTGTCAGTGGCCCTATTGCTGCTG 473
QY 365 AGAAGATGGCCAGGTTGAGCCCAAGATTGGCCCTGACGCAATACGTCAATCTCAAGG 424
DB 474 AGAAGACGAGATACATGCTCAAGCTTGAGCTGATGCACTACATCATCAAGAGG 533
QY 425 ACCTGACCGAAGACCCGAGCCCACTAGCGCGCTATGGAAGAAATTTATGAGATCGCA 484
DB 534 AGCTGACCGAAGACCTGAGCCCACTTGGCGCGCAATGGAGGCCATCTACAACCTCTGGCA 593
QY 485 AGGCAGGTCCATTGGTGTCTCCAACTGGACCATTTGGCGACTTGAGAAGATGTCCAAGT 544
DB 594 GGGCTAAGGCTATGGTGTGTCTCAACTGGACCGCTCAAGACCNCTCAAGCTNCTAGCT 653
QY 545 TCGCCAAAGTTCATGCTTCCAC -GCCAACGATCGAGATTCAACCCCTTCTGCCCAACGAG 603
DB 654 TCGCAAGGTCAAGNCCCGAGTCAACCAATCGAGATCCANCCCTTCTCTCTAACGCC 713
QY 604 GAGTGGTGCAGTACTGCTTCTCAAGACATTTATGCCGTGGCTACTCTCTCTGGGC 663
DB 714 GAGCTCGNCAAGTACTGCGAGGACCGANATTTCTTCCAGCCGGCTACTTGGCGGTTGGG 773
QY 664 TCGCAGAACCAAGTTTCCCAACCGGTTGAGCGGCTCAGCGAGAACAAAGACTCTCTGAACGAG 723

DB 774 TTCCANAACAGTCCCA-CACAGCGAGAGGTCGCAACCAACCCAGCTC--AACGAG 830
QY 724 ATCGCGGAGAGGGCGGAACACCTTCTGCTCAGGTTCTTATTGCTCTGGGGTCT 776
DB 831 GTTGTGAGCGCANGCGCCACGATCTCGACAGTCTGCTGATGGGGCTTT 883

RESULT 2

CF867536 734 bp mRNA linear EST 31-OCT-2003
LOCUS tricol1xm03.b1 T.reesei mycelial culture, Version 6 October 2003
DEFINITION Hypocrea jecorina cDNA clone tricol1xm03, mRNA sequence.
ACCESSION CF867536
VERSION 1
KEYWORDS EST.
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

REFERENCE 1 (bases 1 to 734)
Diener S.E., Dankmeyer L., Dunn-Coleman N., Houfek T.D.,
Mitchell T.K., van Solingen P., Teunissen P.J.M., Ward M. and
Dean R.A.
Analysis of the protein processing and secretion pathways in a
Trichoderma reesei EST dataset
Unpublished (2003)
CONTACT: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph_dean@ncsu.edu
Seq primer: LT-F1 primer.
Location/Qualifiers

1. 734
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tricol1xm03"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pRBP3Y; Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

FEATURES
source

1. 734
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tricol1xm03"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pRBP3Y; Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

ORIGIN

Query Match 27.1%; Score 265; DB 14; Length 734;
Best Local Similarity 66.3%; Pred. No. 9.5e-48;
Matches 379; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
QY 13 AGACTTTCACATTGAGCAACGCGTCAAGATTCCTGGCGTGGTGGTACCTTCGCT 72
DB 163 AAGAGCTACACTCTGAACACCGGTGCGCAAGATTCGCGGGTTCGGACATTCGCC 222
QY 73 AGTGAAGGTTCCAAGGGGAGACCTATACTGTGTCAACACTGCGCCCTGAAGACCGGTTAC 132
DB 223 AATGAGGTGCGCAAGGGCGAGACATACGAGCTGTTACAAAGGCACTGACGTTGATAC 282
QY 133 CGTCACTTGGACTGTGCTGGTACTACCTGACAGAGGTGAGTGGTGGAGGTATCCGT 192
DB 283 CGCCACCTTGTATGCGCGTGGTGTTCACCAACAGAGATGAGTGGTGACCGGTACGC 342
QY 193 GACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAAGACATCTTCGTCTGACCAAGGTG 252
DB 343 GATTTCTCGCCCGCGCCACCGACGTGAACCGAGAGATCTCTTATTGACCAAGTT 402
QY 253 TGAACACCACTCCACCGTTTATGAGACGTCCTCTCTGGTCAATTGACGACTCCCTCGAAGCT 312
DB 403 TGAACACCACTGATGAGCGAGGACGTCAGTGGAGCGCAAGAACTCTGTCGAAAC 462

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2004, 11:17:03 ; Search time 2523 Seconds
(without alignments)
11575.590 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978

Sequence: 1 atgtctaacggaagacttt.....ccagaacctgtctggtga 978

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthm:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353.2	36.1	886	14 CA764322	CA764322 AF53-Rpf
2	265	27.1	734	14 CF867536	CF867536 trico11xm
3	265	27.1	792	14 CB897574	CB897574 trico11xm
4	220.8	22.6	704	14 CF866105	CF866105 trico005xa

5	220.8	22.6	757	14	CB896064
6	188.6	19.3	390	12	BI200418
7	176	18.0	1061	12	BI948347
8	166	17.0	1087	29	CNS06060
9	164.8	16.9	534	9	AA786853
10	160.4	16.4	769	14	CF869025
11	160.4	16.4	823	14	CB899137
12	152.2	15.6	189	9	AA785751
13	149	15.2	779	9	AA263299
14	148.8	15.2	638	13	CA015279
15	148.6	15.2	758	9	AA697426
16	139.4	14.3	856	12	BI948257
17	138.6	14.2	449	12	BM872415
18	136.6	14.0	997	11	CNS08Q80
19	132.2	13.5	827	13	BQ751060
20	131	13.4	1036	11	CNS08Q8N
21	129.4	13.2	955	11	CNS0945X
22	129.4	13.2	1311	11	AY105383
23	128	13.1	643	13	BU645057
24	128	13.1	1021	29	CC594922
25	127.8	13.1	1021	11	CNS094SX
26	127	13.0	790	14	CF642818
27	124.4	12.7	706	14	CF644498
28	122.2	12.5	770	13	EX607579
29	120.2	12.3	1201	13	EX417080
30	119.8	12.2	668	12	BI619770
31	119.6	12.2	1201	13	EX417343
32	119.4	12.2	644	10	BE974991
33	119.4	12.2	665	12	BI623282
34	119.4	12.2	666	12	BI485421
35	119.2	12.2	1019	13	EX331589
36	119.2	12.2	1118	13	EX397710
37	119.2	12.2	1179	13	EX377469
38	118.4	12.1	639	9	AI114144
39	118	12.1	563	12	BI369569
40	117.8	12.0	1044	12	BM561410
41	117.4	12.0	610	12	BG368501
42	117.2	12.0	1017	29	CNS0784Q
43	116.8	11.9	809	12	BI857430
44	116.6	11.9	1176	9	AL547253
45	116.4	11.9	640	9	AI135240

ALIGNMENTS

RESULT 1

CA764322

LOCUS

DEFINITION

CA764322 886 bp mRNA linear EST 08-JAN-2003
AF53-Rpf 03 B08 T7 018.abi IRR1 Drought Stress Panicle Library
Oryza sativa (indica cultivar-group) cDNA clone C000800 5' similar
to Alcohol dehydrogenase [NADP+] (EC 1.1.1.2) (Aldehyde reductase)
(3-DG-reducing enzyme), mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA764322 2 GI:27546321

EST.

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzoideae; Oryza.

1 (bases 1 to 886)

Bennett, J., Arumugam, K., Lafitte, R., Wen, J., Rudd, S. and

Bruskiewicz, R.N.

IRRI Drought Stress Panicle cDNA Library

Unpublished (2002)

On Dec 2, 2002 this sequence version replaced gi:25933577.

Contact: Richard Bruskiewicz

Bionetrics and Bioinformatics Unit

International Rice Research Institute

DAPO 7777, Metro Manila, Philippines

Tel: +63-2-845-0563

Fax: +63-2-845-0606

QY 623 TCTCAAGAACATTATGCCCCGCTGCTACTCTCTCTGGGTGCGAGAA---CCAGGTTTC 679
Db 668 AGTCAAGGACATTACATCACTGCTACAGTCCCTTGGGATCTCCCAACCGGCCATGGG 727
QY 680 CCACACCGGTGACGGGTGAGC---GAGAACAGACTCTGACGAGATCGCGAGAGG 736
Db 728 CCAAGGTGGTGGTCCGGTCACTCTAGAGGAGGCTAAGATCAAGGAAATTCGCGCTAAGA 787
QY 737 GCGGCAACACCCCTTCTCAGGTTCTTATTGCTGGGTCTGCGCGCTGAGTCTGCTTC 796
Db 788 AGAAGAGACCCCTGGACAGATCTTATTCGATACCAAGGTTGAGCGTGCCCAACATTGTTA 847
QY 797 TCCCAAGAGCTCAACCCCAAGCGCATGAGTCCCACTTCAAGAGCATTCAGCTC 852
Db 848 TCCCAAAATCTGTGACCAAGGACCGCATCGAGTCCAACTTCCAGGTTTCGACTTC 903

RESULT 15

ABZ56044

ID ABZ56044 standard; cDNA; 403 BP.

XX AC ABZ56044;

XX DT 28-MAR-2003 (first entry)

XX DE Aspergillus oryzae polynucleotide SEQ ID NO 5157.

XX KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
KW expressed sequence tag; gene; ss.

XX OS Aspergillus oryzae.

XX PN WO200279476-A1.

XX PD 10-OCT-2002.

XX PF 22-MAR-2002; 2002WO-IB000890.

XX PR 30-MAR-2001; 2001JP-00098371.

XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PA (NARE-) NAT RES INST BREWING.

XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.

XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;

PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;

XX WPI; 2003-046817/04.

XX PT Detection of expression of specific Aspergillus genes for monitoring the
PT fermentation and growth conditions of the fungus, using DNA probes.

XX PS Claim 1; SEQ ID NO 5157; 48pp + Sequence Listing; Japanese.

XX CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ5088-ABZ5693), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridising
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 403 BP; 103 A; 107 C; 101 G; 92 T; 0 U; 0 Other;

Query Match 18.0%; Score 176.4; DB 7; Length 403;

Best Local Similarity 70.9%; Pred. No. 3e-37;

Matches 248; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

QY 1 ATGTCTAAACGGAAAGACTTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGGTCCGGCTTT 60
Db 54 ATGTCTCTGGCGTATCCTTCAAAACAAGCAACGGCGTCACCATCCAGGGCGTGGCTTC 113
QY 61 GGTACCTTCGCTAGTGAAGGTTCGAAGGGCGAGACCTATCTGCTGTACCACTCCCTG 120
Db 114 GGTACCTTCGCTAAGCAAGGCTCCAAAGGGCGAGACGTACAAGGCTGTCCGTCACCGTTG 173
QY 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGGTACTTACCTGAACGAGGGTGAGGTTGGT 180
Db 174 AAAGTCGGCTACCGCCACCTCGAATGTGCTGTTCTATGCCAATGAGGACGAAGTAGGA 233
QY 181 GAGGGTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
Db 234 CAGGGGATTCGAAGATTTCCTGAAGAGAGAACCCCGTCGGTTAAACGGGAGGATCTGTTGTG 293
QY 241 TGCACCAAGGTGTGGAACCCACCTCCACCGTTATGAGGACGTCCTCTGGTCCATT-GACGA 299
Db 294 ACGACGAAGGTGTGGAACCCACCTGCATCGGTATGAGATGTTCTCTGGTCCGTTGGAAGA 353
QY 300 CTCCTGGAAGCGTCTTGGACTTGACTACGTTGATATGTTCTCTGTTCACT 349
Db 354 TTCAATTGAAGAACTTGCATTCGAGCTACGCTGATCTGTTTCTCGTGCACT 403

Search completed: May 29, 2004, 11:11:56

Job time : 455 secs

XX Designing synthetic nucleic acid sequences for improved amplification,
 PT expression in host cell, by comparing free energy of folding of a
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement.
 XX
 XX Example 9; Page 62-65; 117pp; English.
 XX
 CC The present sequence encodes a NADPH-dependent aldehyde reductase (ALR).
 CC The polynucleotide sequence was modified using the method of the
 CC invention. The specification describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with a
 CC different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotide, and comparing
 CC this with that of the original polynucleotide. The method is useful for
 CC developing nucleic acid sequences that enhance expression of the encoded
 CC protein in a heterologous host. The design and preparation of the
 CC synthetic genes are used in application of gene shuffling, directed
 CC evolution and molecular breeding methods. The method allows expression of
 CC genes from various organisms such as mammals, plants, yeast, fungi and
 CC bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic
 CC hosts at commercially viable levels, in particular proteins with low
 CC yield such as methionine gamma-lyase from *P. putida*. (Updated on 11-SEP-
 CC 2003 to standardise OS field)
 XX
 XX Sequence 972 BP; 203 A; 337 C; 265 G; 167 T; 0 U; 0 Other;
 SQ
 Query Match 20.4%; Score 199.6; DB 5; Length 972;
 Best Local Similarity 55.5%; Pred. No. 2.5e-43;
 Matches 463; Conservative 0; Mismatches 344; Indels 27; Gaps 3;
 QY 87 GGGCGAGACCTATCTGCTGTACCACTGCGCTCAAGACCGGTACCGTCACTGGACTG 146
 Db |||||
 QY 75 GGGCGAGGTGGCGCAGGCGGTCAAGTGGCGCTGAGTGGATACCGTCACTCGACCT 134
 Db |||||
 QY 147 TGCCTGTGTACTACCTGAACGAGGCTGAGGTGTGGTGGGTATCGGTGATCTCTGGAAGGA 206
 Db |||||
 QY 135 TGCCAAAGTCTACTCGAACCAACCTGAGTGTGGTGGCGCATCA-----A 179
 QY 207 GAACCCCTCGGTGAAGGTGAGGACATCTTGTCTGCACCAAGGTGTGGAAACCACTCCA 266
 Db |||||
 QY 180 GGAGGCTGGCGTCAAGCGCGAGGACCTCTTATCACCTCGAAGCTCTGGAACAACCTGCA 239
 QY 267 CCGTATGAGGACGCTCTCTGCTGCAATGAGCTCCCTGAAGCGTCTTGACTTGA 326
 Db |||||
 QY 240 CCGCGCGAGCAGTGGAGCTGCGCTGCTGAGCACCTCAAGAGCTCGGCTCGAGTA 299
 QY 327 CGTTGATATGTTCTGTTCTACTGGCCCATTTGTCGCGAGAAGAAATGGCCAGGTTAGCC 386
 Db |||||
 QY 300 CCTGACCTTTACCTCATCTGCGCGCTGCGCTTCCCGCGGAGGCGACATCACCA 359
 QY 387 CAAGATTGGCCCTGACGCGCAATACGTCATCTCAAG--GACCTGACCGAGAACCCCGA 443
 Db |||||
 QY 360 GAACTCTTCCGGAAGCCACGAAAGGAGTCAAGCTGACCTGAGGTGAGTCAAGTTCGT 419
 QY 444 GCCCAATGCGCGCTATGGAAGAGATTATGAGATGCGAAGCCAGGTCATTTGGTGT 503
 Db |||||
 QY 420 CGACACGTGGAAGCGATGGTCAAGCTTCTCGACACTGGCAAGGTCAAGCGATCGGCGT 479
 QY 504 CTCCAACTGGACCATTCGCGACCTTGAAGATGTCACAGTTCGCGAAGTCACTGCTCA 563
 Db |||||
 QY 480 TTCCAATTCGACCGGAAGATGTTGAGCGCATCATGAGGCTACCGCGTGAACCCCTC 539
 QY 564 CGCCAAACAGATCAGATTTACACCCCTTCTCCGCCCAAGAGGAGTGTGTGAGTACTGCTT 623
 Db |||||
 QY 540 CGTCAACAGATCGAGCGTCAACCTCTCTTCTCCAGCGAGTCTATCGCCACCAACAA 599
 QY 624 CTCCAAGACATTTATGCCCGTGGCTACTCTCTCTGCGCTCGGAGAACAGGTTCACAC 683
 Db |||||
 QY 600 GGCCAAAGAACATTACATTTACCGCATATCTCTCTCTCGGTAAACAAACCGTGGCGCGCC 659

QY 684 CACCGGTGAGCGGTGTCAGCGAGAACAGAGACTTGTGAACGAGATGCGCGAGAGGGCGCAA 743
 Db |||||
 QY 660 TCTTCTTGTCCAGCACCGCGAGATCAAGCGC-----ATCGCGAGAGAAACGCGTG 710
 Db |||||
 QY 744 CACCTTCTGCTCAGTCTCTTATTTGCTGGGTCTGCGCGTGGTACGTCGTTCTCCCAA 803
 Db |||||
 QY 711 CAGCGCGCTCAGGTCTCTCAITTCCTGGGCACTGTTGGCGGCACTCGGTATATCCCAA 770
 Db |||||
 QY 804 GAGCTCCAAACCCCAAGCGCATTTGAGTCCAACTTCAAGAGCATTTGAGTCTCCGATGCCGA 863
 Db |||||
 QY 771 GTCGCTCAACCCCTCCCGCATTTGGGAGAACTTCAAGCAGGTCTGCTCTCGCAGGAGA 830
 QY 864 CTTTGAAGCATCAATGCGGTGCCAAGGTCGTCACCTTCGTTCTGTCACAT 917
 Db |||||
 QY 831 CGTGATGCGGTCAGCAAGCTCGCGAGGTTGCGGCGCGAGGCGCTACAACAT 884
 RESULT 13
 AAT29159
 ID AAT29159 standard; cDNA to mRNA; 1055 BP.
 XX
 AC AAT29159;
 XX
 DT 16-OCT-2003 (revised)
 DT 18-NOV-1996 (first entry)
 XX
 DE Carboxyl reductase gene.
 XX
 KW carbonyl reductase; R-gamma-substituted-beta-hydroxybutyrate;
 KW gamma-substituted acetate; raw material; drug synthesis;
 KW agricultural chemical production; ds.
 XX
 OS Sporidibolus salmonicolor; IP01038.
 XX
 FH Key Location/Qualifiers
 FT CDS 30..1001
 FT /*tag= a
 FT /product= "carbonyl_reductase"
 FT mat_peptide 33..999
 FT /*tag= b
 XX
 XX JP08103269-A.
 XX
 XX 23-APR-1996.
 XX
 XX 07-OCT-1994; 94JP-00244090.
 XX
 XX 07-OCT-1994; 94JP-00244090.
 XX
 PA (ELED) DENKI KAGAKU KOGYO KK.
 XX
 XX WPI; 1996-253873/26.
 XX
 XX P-PSDB; AAR96294.
 XX
 PT Carboxyl reductase gene and protein - used in the production of R-gamma-
 PT substituted-beta-hydroxy-butyrate, for use in drugs and agrochemicals.
 XX
 PS Claim 1; Page 9-10; 14pp; Japanese.
 XX
 CC The present sequence encodes a carbonyl reductase (ALR) isolated from
 CC *Sporidibolus salmonicolor* IP01038. The enzyme is used in the synthesis
 CC of an (R)-gamma-substituted-beta-hydroxybutyrate (GSBH) from a gamma-
 CC substituted acetate. The GSBH is useful as a raw material for the
 CC synthesis of drugs and agricultural chemicals. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 1055 BP; 224 A; 365 C; 284 G; 182 T; 0 U; 0 Other;
 Query Match 20.4%; Score 199.6; DB 2; Length 1055;
 Best Local Similarity 55.5%; Pred. No. 2.5e-43;
 Matches 463; Conservative 0; Mismatches 344; Indels 27; Gaps 3;
 QY 87 GGGCGAGACCTATCTGCTGTACCACTGCGCTCAAGACCGGTACCGTCACTGGACTG 146

AC AAT29160;
 XX 16-OCT-2003 (revised)
 DT 18-NOV-1996 (first entry)
 XX
 DE Carbonyl reductase coding sequence.
 XX carbonyl reductase; R-gamma-substituted-beta-hydroxybutyrate;
 KW gamma-substituted acetate; raw material; drug synthesis;
 KW agricultural chemical production; ds.
 XX
 OS Sporidiobolus salmonicolor; IF01038.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..972
 FT /tag= a
 FT /product= "carbonyl_reductase"
 FT mat_peptide 4..969
 FT /tag= b
 XX
 FN JP08103269-A.
 XX
 XX 23-APR-1996.
 PD
 XX 07-OCT-1994; 94JP-00244090.
 PF
 XX 07-OCT-1994; 94JP-00244090.
 PR
 XX (ELED) DENKI KAGAKU KOGYO KK.
 PA
 XX WPI; 1996-253873/26.
 DR
 XX P-PSDB; AAR96294.
 XX
 FT Carbonyl reductase gene and protein - used in the production of R-gamma-
 FT substituted-beta-hydroxybutyrate, for use in drugs and agrochemicals.
 XX
 PS Claim 3; Page 10-11; 14pp; Japanese.
 XX
 CC The present sequence encodes a carbonyl reductase (ALD) isolated from
 CC Sporidiobolus salmonicolor IF01038. The enzyme is used in the synthesis
 CC of an (R)-gamma-substituted-beta-hydroxybutyrate (GSBH) from a gamma-
 CC substituted acetate. The GSBH is useful as a raw material for the
 CC synthesis of drugs and agricultural chemicals. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 972 BP; 203 A; 337 C; 265 G; 167 T; 0 U; 0 Other;

Query Match 20.4%; Score 199.6; DB 2; Length 972;
 Best Local Similarity 55.5%; Pred. No. 2.5e-43;
 Matches 463; Conservative 0; Mismatches 344; Indels 27; Gaps 3;

Qy 87 GGGGAGACCTATCTACTGTCTGACCACTGCCCTGACAGCCGGTTACCGTCACCTTGACATG 146
 |||||
 Db 75 GGGGAGGTGGGCGAGGGGCTCAGGTGCGCATCGAGACTGGATACCGTCACCTCGACCT 134
 |||||
 Qy 147 TGCTGTGTACTCTGAAAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 206
 |||||
 Db 135 TGCCAAAGGTCTACTCGAACCAACCTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 179
 |||||
 Qy 207 GAACCCCTCGTGAAGGTGAGGACATCTGCTGTCACCAAGGTGGAACCACTCCA 266
 |||||
 Db 180 GGAGGTGGGTGAAGCGGAGGACCTCTCATCACCCTCGAGCTCTGGAGCAACTCGCA 239
 |||||
 Qy 267 CCGTTATGAGGACCTCTCTGGTCCATTGACGACTCCCTGAAGCGTCTTGGAATTGACTA 326
 |||||
 Db 240 CCGCCCGAGCGGTGAGGCTGCGCTGACGACACCCCTCAAGGAGCTCGGCTCGAGTA 299
 |||||
 Qy 327 CGTTGATCTCTCTGCTTCACTGGCCATTGTCGCGAGAGAAATGGCCAGGCTGAGCC 386
 |||||
 Db 300 CCTCGACCTTTACCTCATTTACTGGCCGCTGCGTTCCCGCCGAGGCGACATCACCCA 359
 |||||
 Qy 387 CAAGATTGGCCCTGACGCGCAATACGTATTCTCAAG--GACCTGACCGAGAACCCCGA 443
 |||||

Db 360 GAACCTCTTCCGAGAGGCCAAGCAAGAGAGGTCAAGCTCGACTCGAGGTCAAGCTCGT 419
 Qy 444 GCCACATGGCGCGCTATGGAGAAATTATGAGATCGCAAGGCCAGGTCCATTGGTGT 503
 |||||
 Db 420 CGACACGTGGAAAGCGGATGGTCAAGCTTCTCGACACTGGCAAGGTCAAGGCGATCGGCGT 479
 |||||
 Qy 504 CTCGAACCTGGACCATTTGCCGACCTTGAGAAGATGTCGAAGTTCGAAGTTCGCAAGTCAATGCCCTCA 563
 |||||
 Db 480 TTCCAACTTCGACGCGAAGATGGTCGAGCCCATCATCGAGGCTACCGGCGTGACCCCTC 539
 |||||
 Qy 564 GCGCAACCAAGATCGAGATTACCCCTTCTCTGCCAAACGAGGAGTGTGTCAGTACTGTGTT 623
 |||||
 Db 540 CGTCAACCAAGATCGAGCGTCAACCTCTCTTCTCCAGCCGAGCTCATCGCCACCAAA 599
 |||||
 Qy 624 CTCGAAGAAATATGCCCCGTGGCTACTCTCTCTGCGCTCGAGAACCAAGGTTCAC 683
 |||||
 Db 600 GCGCAAGAAATATTCACATTACCGCATACTCTCTCTCGGTAAACAACACCGTCGCGGCGCC 659
 |||||
 Qy 684 CACCGGTGAGCGGTGTCAGCGAGAAACAAGACTCTGAACGAGATCGCCGAGAGGCGGCA 743
 |||||
 Db 660 TCTTCTTGTCCAGCACCCGAGATCAAGCGC-----ATCGCGAGAGAACGCTG 710
 |||||
 Qy 744 CACCTTGTCTCAGGTCTTATGTGCTGGGTCTCGCGGTGGCTAGTCTCTCTCTCCCAA 803
 |||||
 Db 711 CACGCCGCTCAGGTCTCTCATTGCTGGCCCATCGTTGGCGGCCACTCGGTATATCCCAA 770
 |||||
 Qy 804 GAGCTCCAAACCCAGCGCATTTGAGTCCAACTTCAAGAGCATTTGAGCTCTCCGATCGCA 863
 |||||
 Db 771 GTGCGTCAACCCCTCCCGCATTTGGGAGAACTTCAAGCAGTCTCGCTCTCGAGGAGA 830
 |||||
 Qy 864 CTTTGAAGCCATCAATGCGTTGCCAAGGGTGTGTCACCTCCGTTTCGTTCAACAT 917
 |||||
 Db 831 CGTCGATCCGCTCAGCAAGCTCGCGAGGTTGCGGCGCAGGCGCTACAACAT 884
 |||||

RESULT 12
 AAH74577
 ID AAH74577 standard; DNA; 972 BP.
 XX
 AC AAH74577;
 XX
 DT 11-SEP-2003 (revised)
 DT 15-OCT-2001 (first entry)
 XX
 DE Nucleotide sequence of a NADPH-dependent aldehyde reductase (ALR).
 XX
 XX Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding;
 KW NADPH-dependent aldehyde reductase; ss.
 XX
 OS Sporidiobolus salmonicolor.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..972
 FT /tag= a
 FT /transl_except= (pos: 316..321, aa: Ile)
 FT /product= "NADPH-dependent aldehyde reductase (ALR)"
 XX
 PN WO200155342-A2.
 XX
 XX 02-AUG-2001.
 PD
 XX 31-JAN-2001; 2001WO-US003186.
 PF
 XX 31-JAN-2000; 2000US-00494921.
 PR 08-DEC-2000; 2000US-00734237.
 PR
 XX (BIOC-) BIOCATALYTICS INC.
 PA
 XX Rozzell DJ, Bui P, Hua L;
 PI
 XX WPI; 2001-483235/52.
 DR
 DR P-PSDB; AAG63561.

XX PD 12-JUN-2002.
XX PF 07-DEC-2001; 2001EP-00310251.
XX KW 07-DEC-2000; 2000JP-00372704.
XX KW 15-JAN-2001; 2001JP-00006144.
XX PR 02-FEB-2001; 2001JP-00026594.
XX PR 11-JUN-2001; 2001JP-00175175.
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;
XX PF WPI; 2002-550350/59.
XX DR
XX PT New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by
XX PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals
XX PT and agrochemicals.
XX PS Example 1; Page 38; 56pp; English.
XX CC The present sequence represents a PCR amplified fragment of a gene
XX CC encoding a Penicillium citrinum protein. This protein is capable of
XX CC producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-
XX CC bromo-3-oxobutanoate. The amplified fragment was inserted into plasmid
XX CC pBR-2. The protein and polynucleotides are useful for producing optically
XX CC active (S)-4-bromo-3-hydroxybutanoate, which is useful as an intermediate
XX CC in the production of pharmaceuticals and agrochemicals. The (S)-4-bromo-3-
XX CC -hydroxybutanoate produced can also be used to produce 4-cyano-3-
XX CC hydroxybutanoic acid
XX SQ Sequence 743 BP; 183 A; 202 C; 156 G; 167 T; 0 U; 35 Other;
XX
XX Query Match 42.1%; Score 412; DB 6; Length 743;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-100;
XX Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 567 CAACGATCGAGATTACCCCTTCCTGCCAACAGGAGTGCTGCAGTACTGCTTC 626
Db 1 CAACGATCGAGATTACCCCTTCCTGCCAACAGGAGTGCTGCAGTACTGCTTC 60
QY 627 CAAGAACATTATGCGGTGGCTACTCTCTCTGGCTCGAGAACAGGTTCACCCAC 686
Db 61 CAAGAACATTATGCGGTGGCTACTCTCTCTGGCTCGAGAACAGGTTCACCCAC 120
QY 687 CGGTGAGCGGTGAGCGAGAACAGACTCTCAACGAGATCGCGAGAGGCGGCAACAC 746
Db 121 CGGTGAGCGGTGAGCGAGAACAGACTCTCAACGAGATCGCGAGAGGCGGCAACAC 180
QY 747 CTTGCTCAGGTTCTTATTGCTGGGTCTGCGCGGTGCTGCTGCTTCTCCCAAGAG 806
Db 181 CTTGCTCAGGTTCTTATTGCTGGGTCTGCGCGGTGCTGCTGCTTCTCCCAAGAG 240
QY 807 CTCACACCCCAAGCGCATTTAGTCCACTTCAAGAGCATTTAGTCTCCGATCCGACTT 866
Db 241 CTCACACCCCAAGCGCATTTAGTCCACTTCAAGAGCATTTAGTCTCCGATCCGACTT 300
QY 867 TGAAGCCATCAATGCCGTGTCAGAGGTCGTCACTTCGGTTTCGTCAACATGAAGATAC 926
Db 301 TGAAGCCATCAATGCCGTGTCAGAGGTCGTCACTTCGGTTTCGTCAACATGAAGATAC 360
QY 927 TTTCGGATATCATCTGCGCCGAGGAGACCGCCCAAGAACCTGTCGCTGA 978
Db 361 TTTCGGATATCATCTGCGCCGAGGAGACCGCCCAAGAACCTGTCGCTGA 412

RESULT 10
ABL59387
ID ABL59387 standard; DNA; 331 BP.
XX
AC ABL59387;
XX

DT 22-OCT-2002 (first entry)
XX DE PCR amplicon of a protein producing (S)-4-bromo-3-hydroxybutanoate.
XX KW (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
XX KW agrochemical; 4-cyano-3-hydroxybutanoic acid; ss.
XX OS Penicillium citrinum.
XX PN EP1213354-A2.
XX PD 12-JUN-2002.
XX PF 07-DEC-2001; 2001EP-00310251.
XX PR 07-DEC-2000; 2000JP-00372704.
XX PR 15-JAN-2001; 2001JP-00006144.
XX PR 02-FEB-2001; 2001JP-00026594.
XX PR 11-JUN-2001; 2001JP-00175175.
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;
XX PF WPI; 2002-550350/59.
XX DR
XX PT New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by
XX PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals
XX PT and agrochemicals.
XX PS Example 1; Page 38; 56pp; English.
XX CC The present sequence represents a PCR amplified fragment of a gene
XX CC encoding a Penicillium citrinum protein. This protein is capable of
XX CC producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-
XX CC bromo-3-oxobutanoate. The amplified fragment was inserted into plasmid
XX CC pBR-1. The protein and polynucleotides are useful for producing optically
XX CC active (S)-4-bromo-3-hydroxybutanoate, which is useful as an intermediate
XX CC in the production of pharmaceuticals and agrochemicals. The (S)-4-bromo-3-
XX CC -hydroxybutanoate produced can also be used to produce 4-cyano-3-
XX CC hydroxybutanoic acid
XX SQ Sequence 331 BP; 70 A; 93 C; 92 G; 75 T; 0 U; 1 Other;
XX
XX Query Match 27.1%; Score 265; DB 6; Length 331;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-61;
XX Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTAACGGAAGACTTTCACATTGAGCAACGCGTCACAGATTCCTGGCGTCGGCTTT 60
Db 67 ATGCTAACGGAAGACTTTCACATTGAGCAACGCGTCACAGATTCCTGGCGTCGGCTTT 126
QY 61 GGTACTCTCGCTAGTGAAGGTTCCAAAGGCGAGACCTATCTGCTGCACCACTGCCCTG 120
Db 127 GGTACTCTCGCTAGTGAAGGTTCCAAAGGCGAGACCTATCTGCTGCACCACTGCCCTG 186
QY 121 AAGACGGTTACCGTCACCTTGACCTGTGCTGGTACTACTCTGAAAGAGGTTGGT 180
Db 187 AAGACGGTTACCGTCACCTTGACCTGTGCTGGTACTACTCTGAAAGAGGTTGGT 246
QY 181 GAGGTTATCCGTGACTTCTCTGAAAGAGAACCCCTCGGTGAAGCGTGAAGACATCTTCCTC 240
Db 247 GAGGTTATCCGTGACTTCTCTGAAAGAGAACCCCTCGGTGAAGCGTGAAGACATCTTCCTC 306
QY 241 TCACCAAGGTTGTGAACCACTCC 265
Db 307 TCACCAAGGTTGTGAACCACTCC 331
XX
XX RESULT 11
XX AAT29160
XX ID AAT29160 standard; cDNA to mRNA; 972 BP.
XX

```
Query Match 42.6%; Score 417; DB 6; Length 417;
Best Local Similarity 100.0%; Pred. No. 6.9e-102;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 CCAAGGTGTGGAACACCTCCACGGTTATGAGGACGCTCTGCTGTCATTGAGGACTCCC 304
D 1 CCAAGGTGTGGAACACCTCCACGGTTATGAGGACGCTCTGCTGTCATTGAGGACTCCC 60
QY 305 TGAAGCGCTCTTGGAATTGACTACGTTGATATGTTCTCTGTTCACTGGCCCATTTGTCGG 364
D 61 TGAAGCGCTCTTGGAATTGACTACGTTGATATGTTCTCTGTTCACTGGCCCATTTGTCGG 120
QY 365 AGAAGAAATGCGCAGGGTGAGCCCAAGATTTGGCCCTGACGCAATATGCTCAAG 424
D 121 AGAAGAAATGCGCAGGGTGAGCCCAAGATTTGGCCCTGACGCAATATGCTCAAG 180
QY 425 ACCTGACCGAGAACCCGAGCCACATGCGCGCTATGAGAGAGATTTATGAGGATCGCA 484
D 181 ACCTGACCGAGAACCCGAGCCACATGCGCGCTATGAGAGAGATTTATGAGGATCGCA 240
QY 485 AGGCCAGGTCCATTGGTGTCTCCAACTGGACCATTTGCGGACCTTGAGAGATGTCAAAGT 544
D 241 AGGCCAGGTCCATTGGTGTCTCCAACTGGACCATTTGCGGACCTTGAGAGATGTCAAAGT 300
QY 545 TCGCCAGGTTCATGCTCAGCCCAACAGATCGAGATTCACTCCCTTCCGCAAGGAGG 604
D 301 TCGCCAGGTTCATGCTCAGCCCAACAGATCGAGATTCACTCCCTTCCGCAAGGAGG 360
QY 605 AGCTGGTGCAGTACTGCTTCTCCAGAACATTTATGCCCGTGGCTACTCTCTCTGG 661
D 361 AGCTGGTGCAGTACTGCTTCTCCAGAACATTTATGCCCGTGGCTACTCTCTCTGG 417

RESULT 8
ID ABZ53802 standard; cDNA; 587 BP.
XX
AC ABZ53802;
XX
DT 28-MAR-2003 (first entry)
XX
XX Aspergillus oryzae polynucleotide SEQ ID NO 2915.
DE
XX
XX Aspergillus oryzae; fermentation; fungus; industrial; EST;
KW expressed sequence tag; gene; ss.
XX
XX Aspergillus oryzae.
OS
XX
XX WO200279476-A1.
XX
XX 10-OCT-2002.
XX
XX 22-MAR-2002; 2002WO-IB000890.
XX
XX 30-MAR-2001; 2001JP-00098371.
XX
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX
XX (NARE-) NAT RES INST BREWING.
XX
XX (NORQ) NAT FOOD RES INST MIN AGRIC.
XX
XX Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
XX Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
XX WPI; 2003-046817/04.
XX
XX Detection of expression of specific Aspergillus genes for monitoring the
XX fermentation and growth conditions of the fungus, using DNA probes.
XX
XX Claim 1; SEQ ID NO 2915; 48pp + Sequence Listing; Japanese.
XX
XX The invention relates to a polynucleotide having any of 6006 specific
XX sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
XX specific culture conditions including one or more of eutrophic,
XX
```

```
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridising
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: the
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 587 BP; 142 A; 177 C; 151 G; 117 T; 0 U; 0 Other;
```

```
Query Match 42.5%; Score 415.4; DB 7; Length 587;
Best Local Similarity 83.8%; Pred. No. 2.1e-101;
Matches 470; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
```

```
QY 1 ATGTCTAAGCGAAGAACTTTACATTTGAGCAACGGCGTCAAGATTCCTGGCGTGGCTTT 60
D 27 ATGTCTAAGCGAAGAACTTTACATTTGAGCAACGGCGTCAAGATTCCTGGCGTGGCTTT 86
QY 61 GGTACTTTCGCTAGTAGAAGTTCCAGGGCGAGACTATCTGCTGTCACCACTGCCCTG 120
D 87 GGTACTTTCGCTAGTAGAAGTTCCAGGGCGAGACTATCTGCTGTCACCAAGGCCCTC 146
QY 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGGTACTACTCTGAACGAGGTCAGGTTGGT 180
D 147 GAGACCGGATACCGTCACTTGGACTGTGCTGGTACTACTCTCAACGAGGATGAGTTGGT 206
QY 181 GAGGATATCCGTTGACTTCTGTAAGGAGAACCCCTCGGTGAGCGTGAGGACATCTTCGTC 240
D 207 GATGGTATCCATGACTTCTCAAGAGAACCCCTCCGTCAGCGGGAAGACATCTTCGTC 266
QY 241 TGCACACAGGTGTGGAACACCACTCCACCGTTATGAGGACCTCTCTGCTCAATTGACGAC 300
D 267 TGCACCTAAGTCTGGAATCACCTCCACCGTCCCGAGGACGTCAGTGTGCTGATAC 326
QY 301 TCCCTGAAGCGTCTTGGACTTGGACTTACGTTGATATGTTCTCTGTTTCACTGCGCCATTCCT 360
D 327 TCCCTGAAGAGACTCCGACTGGACTTACGTTGACTTCTTCTTGTCTCACTGCGCCATTCG 386
QY 361 GCGGAGAGATGCGCCAGGTGAGCCCAAGATTGCGCTGACGGCAATAGTCACTTC 420
D 387 TCCGAGAGGAGGACCAAGAGAAACCAAGATTGCGCTGACGGCAATAGTCACTTC 446
QY 421 AAGGACCTTGACCGAGAACCCCGAGCCACATGCGCGCTATGAGAGAGATTTATGAGAT 480
D 447 AAGGAGCTCACTGAGAACCCCGAGCCACATGCGCGCTATGAGAGAGATTTATGAGGAC 506
QY 481 CCGAAGGCGAGTCCATTGGTGTCTCCAACTGGACCATTCGGACCTTGAGAGATGTC 540
D 507 GGCAGAGGCGCAAGGCCATCGGTGTTTCCAACTGGACCATTCGGCGGTCTCGAAGAGCTGTC 566
QY 541 AAGTTTCGCCAAGGTCATGCT 561
D 567 AAGTTTCGCCGAGATCAAGCT 587
```

```
RESULT 9
ID ABL59388 standard; DNA; 743 BP.
XX
XX ABL59388;
XX
XX 22-OCT-2002 (first entry)
XX
XX PCR amplicon of a protein producing (S)-4-bromo-3-hydroxybutanoate.
XX (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
XX agrochemical; 4-cyano-3-hydroxybutanoic acid; ss.
XX
XX Penicillium citrinum.
XX
XX BP1213354-A2.
XX
XX
```

XX PD 28-SEP-2000.
XX PF 22-MAR-2000; 2000WO-US007781.
XX PR 22-MAR-1999; 99US-00273623.
XX PA (NOVO) NOVO NORDISK BIOTECH INC.
XX PA (NOVO) NOVO NORDISK AS.
XX PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX DR WPI; 2000-594572/56.
XX PT Monitoring differential expression of genes in filamentous fungal cells
XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a
XX PT substrate of expressed sequence tags.
XX PS Claim 88; Page 2986; 3161pp; English.
XX CC The present invention describes a method for monitoring differential
XX CC expression of genes in a first filamentous fungal (FF) cell relative to
XX CC expression of the same genes in one or more second filamentous fungal
XX CC cells. The method uses fluorescence-labeled nucleic acids isolated from
XX CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
XX CC are used in the methods for monitoring differential expression of genes
XX CC in a first filamentous fungal (FF) cell relative to expression of the
XX CC same genes in one or more second filamentous fungal cells. Monitoring the
XX CC global expression of genes from FF cells allows the production potential
XX CC of the microorganisms to be improved. New genes may be discovered,
XX CC possible functions of unknown open reading frames can be identified and
XX CC gene copy number variation and stability can be monitored. The expression
XX CC of genes can be used to study how FF cells adapt to changes in culture
XX CC conditions, environmental stress, spore morphogenesis, recombination,
XX CC metabolic or catabolic pathway engineering. Using ESTs provides several
XX CC advantages over genomic or random cDNA clones including elimination of
XX CC redundancy as one spot on an array equals one gene or open reading frame,
XX CC and organisation of the microarrays based on function of the gene
XX CC products to facilitate analysis of the results. AAF07478 to AAF11247
XX CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents
XX CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from
XX CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from
XX CC *Trichoderma reesei*, which are all specifically claimed in the present
XX CC invention.
XX SQ Sequence 637 BP; 162 A; 189 C; 154 G; 132 T; 0 U; 0 Other;
Query Match 43.0%; Score 420.2; DB 3; Length 637;
Best Local Similarity 82.0%; Pred. No. 1.e-102;
Matches 496; Conservative 0; Mismatches 108; Indels 1; Gaps 1;
QY 1 ATGTCTACGGAAGACTTTCACATTTGAGCAACGGCTCAAGATCTCTGGCGTGGCTTT 60
DB 30 ATGTCTACGGAAGACTTTCACATTTGAGCAACGGCTCAAGATCTCTGGCGTGGCTTT 89
QY 61 GGTACCTTCGCTAGTGAAGGTTTCAAGGGCGAGACCTATCTGTGTACCACTGCCCTG 120
DB 90 GGTACCTTCGCTAGTGAAGGTTTCAAGGGCGAGACCTATCTGTGTACCACTGCCCTG 149
QY 121 AAGCCGGTTACCGTCACTTGGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
DB 150 GAGACCGGATACCGTCACTTGGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 209
QY 181 GAGGTATCCGTGACTTCTCAAGGAGAAACCCCTCGGTGAAGCGTGAAGCATCTTCGTC 240
DB 210 GATGGTATCCATGACTTCTCAAGGAGAAACCCCTCGGTGAAGCGTGAAGCATCTTCGTC 269
QY 241 TGCACCAAGGTGTGGAACCACTCCACGGTATGAGGACGTCCTCTGCTCAATGACGAC 300
DB 270 TGCACCAAGGTGTGGAACCACTCCACGGTATGAGGACGTCCTCTGCTCAATGACGAC 329
QY 301 TCCCTGAGCGTCTGGACTTGAATGATGATGATGATGATGATGATGATGATGATGATG 360

DB 330 TCCTGAGAAACTCGACTGACTAGTGTACCTCTTTCTGTCCACTGGCCCATGGCC 389
QY 361 GCGGAGAGATGGCCAGGGTGAGCCCAAGATTGGCCCTCGACGGCAAAATACGTCAATTC 420
DB 390 TCCGAGAGGAGGACCCAGGAAAAACCAAGATTGGCCCTGACCGCAAGTACGTCACTCT 449
QY 421 AAGGACCTGACCGAGAACCCCGAGCCCAATGCGCGCTATGAGGAAGATTATGAGGAT 480
DB 450 AAGGAGCTCAGTGAGAACCCCGAGCCCAATGCGCGCTATGAGGAAGATTATGAGGAT 509
QY 481 CGCAAGGCCAGGTCCTCAATTTGTTCTCTCAACTGGACCATTTGCCGACTTGGAGAAGATGTC 540
DB 510 CGCAAGGCCCAAGCCATCGGTGT-TTCAACTGGACCATTTCCGGGTCTCGAGAAGCTGTT 568
QY 541 AAGTTCCGCAAGCTATGCTTCACTGCTTCAACTGGACCATTTGCCGACTTGGAGAAGATGTC 600
DB 569 AAGTTCCGCGAGATCAAGCCCTTATGTTAACCAAGATGAGATCCACCCCTTTCTGCGCAAC 628
QY 601 GAGGA 605
DB 629 AAGA 633
RESULT 7
ID ABL59391 standard; DNA; 417 BP.
XX ABL59391;
XX AC ABL59391;
XX DT 22-OCT-2002 (first entry)
XX DE PCR amplicon of a protein producing (S)-4-bromo-3-hydroxybutanoate.
XX DE (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
XX KW agrochemical; 4-cyano-3-hydroxybutanoic acid; ss.
XX OS Penicillium citrinum.
XX PN BP1213354-A2.
XX PD 12-JUN-2002.
XX PF 07-DEC-2001; 2001EP-00310251.
XX PR 07-DEC-2000; 2000JP-00372704.
XX PR 15-JAN-2001; 2001JP-00006144.
XX PR 02-FEB-2001; 2001JP-00026594.
XX PR 11-JUN-2001; 2001JP-00175175.
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;
XX DR WPI; 2002-550350/59.
XX PT New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by
XX PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals
XX and agrochemicals.
XX PS Example 1; Page 39; 56pp; English.
XX CC The present sequence represents a PCR amplified fragment of a gene
XX CC encoding a penicillium citrinum protein. This protein is capable of
XX CC producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-
XX CC bromo-3-oxobutanoate. The amplified fragment was inserted into plasmid
XX CC pBR-3. The protein and polynucleotides are useful for producing optically
XX CC active (S)-4-bromo-3-hydroxybutanoate, which is useful as an intermediate
XX CC in the production of pharmaceuticals and agrochemicals. The (S)-4-bromo-3-
XX CC hydroxybutanoate produced can also be used to produce 4-cyano-3-
XX CC hydroxybutanoic acid
XX SQ Sequence 417 BP; 94 A; 126 C; 105 G; 92 T; 0 U; 0 Other;

Db 601 GAGGAGCTGGTGCAGTACTGCTTCTCCAGAACATTATATGCCCGTGGCTACTCTCTCTG 660
Qy 661 GGCTCGCAGAACACGGTTCACACACCGGTGAGGGGTGACGAGACAGACTCTGAC 720
Db 661 GGCTCGCAGAACACGGTTCACACACCGGTGAGGGGTGACGAGACAGACTCTGAC 720
Qy 721 GAGATCCCGCAGAACGGCGGCAACACCTTGTCTCAGGTTCTTATGCTGGGCTCTGCGC 780
Db 721 GAGATCCCGCAGAACGGCGGCAACACCTTGTCTCAGGTTCTTATGCTGGGCTCTGCGC 780
Qy 781 CGTGCTACGTGCTTCTCCCAAGAGCTCCAAACCCCAAGCGCATGAGTCCCACTTCAAG 840
Db 781 CGTGCTACGTGCTTCTCCCAAGAGCTCCAAACCCCAAGCGCATGAGTCCCACTTCAAG 840
Qy 841 AGCATTCAGCTCTCCGATGCGGACTTTGAAGCCATCAATGCCGTGTCGAAGGGTCTGTCAC 900
Db 841 AGCATTCAGCTCTCCGATGCGGACTTTGAAGCCATCAATGCCGTGTCGAAGGGTCTGTCAC 900
Qy 901 TTCGTTTCGTCACATGAAGGATACCTTCGATATGATCTGCGCCGAGGAGACCGCC 960
Db 901 TTCGTTTCGTCACATGAAGGATACCTTCGATATGATCTGCGCCGAGGAGACCGCC 960
Qy 961 AAGAACCTGCTGCGTGA 978
Db 961 AAGAACCTGCTGCGTGA 978

RESULT 5

ABL59384
ID ABL59384 standard; DNA; 697 BP.

XX AC ABL59384;

DT 22-OCT-2002 (first entry)

DE PCR amplicon of a protein producing (S)-4-bromo-3-hydroxybutanoate.

XX (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
KW agrochemical; 4-cyano-3-hydroxybutanoic acid; ss.

XX Penicillium citrinum.

XX EP1213354-A2.

XX PD 12-JUN-2002.

XX PF 07-DEC-2001; 2001EP-00310251.

XX PR 07-DEC-2000; 2000JP-00372704.

XX PR 15-JAN-2001; 2001JP-00006144.

XX PR 02-FEB-2001; 2001JP-00026594.

XX PR 11-JUN-2001; 2001JP-00175175.

XX PA (SUMO) SUMITOMO CHEM CO LTD.

XX PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;

XX DR WPI; 2002-550350/59.

XX New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by
PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals
PT and agrochemicals.

XX Example 1; Page 37; 56pp; English.

XX The present sequence represents a PCR amplified fragment of a gene
CC encoding a Penicillium citrinum protein. This protein is capable of
CC producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-
CC bromo-3-oxobutanoate. The amplified fragment was inserted into plasmid
CC p27-1. The protein and polynucleotides are useful for producing optically
CC active (S)-4-bromo-3-hydroxybutanoate, which is useful as an intermediate
CC in the production of pharmaceuticals and agrochemicals. The (S)-4-bromo-3-

CC -hydroxybutanoate produced can also be used to produce 4-cyano-3-
XX hydroxybutanoic acid

SQ Sequence 697 BP; 161 A; 194 C; 172 G; 153 T; 0 U; 17 Other;

Query Match 60.2%; Score 588.6; DB 6; Length 697;

Best Local Similarity 94.4%; Pred.No. 6.5e-148;

Matches 611; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Qy 4 TCTAACGGAAGACATTTTCACATTCAGCAACGGCGTCAAGATTCCTGGCGTGGTTTGGT 63

Db 52 TCCACGGAANACTTTTCACACTGAGCAACGGCGTCAAAATTCCTGGCGTGGTTTGGT 111

Qy 64 ACCTTCGCTAGTGAAGGTTTCAAGGGCGAGACCTATCTGCTGTCTGTCACACTGCCCTGAAG 123

Db 112 ACCTTCGCTAGTGAAGGTTTCAAGGGCGAACCCTATCTGCTGTCTGTCACACTGCCCTGAAG 171

Qy 124 ACCGGTTACCGTCACTTGGACTGTGCTGTGTACTACCTGAACGAGGGTGAAGTTGGTGAG 183

Db 172 ACCGGTTACCGTCACTTGGACTGTGCTGTGTACTACCTGAACGAGGGTGAAGTTGGTGAG 231

Qy 184 GGTATCCGTCGCTTCTCTGAAGAGAACCCCTCGGTGAAGCGTGAAGCATCTTCTGCTGCG 243

Db 232 GGTATCCGTCGCTTCTCTGAAGAGAACCCCTCGGTGAAGCGTGAAGCATCTTCTGCTGCG 291

Qy 244 ACCAAGGTTGGACCACTCCACCGTTATGAGGACGTCCTCTGCTGTCATTGAACACTCC 303

Db 292 ACCAAGGTTGGACCACTCCACCGTTATGAGGACGTCCTCTGCTGTCATTGAACACTCC 351

Qy 304 CTGAAGCGTCTTGGACTTGGACTAGCTTGCATATGTTCTGCTTCACTGCGCCATTGCTGCC 363

Db 352 CTGAAGCGTCTTGGACTTGGACTAGCTTGCATATGTTCTGCTTCACTGCGCCATTGCTGCC 411

Qy 364 GAGAAGATGGCCAGGTTGAGCCCAAGATTCGCTGAGCGCAATACGTCTCTCAAG 423

Db 412 GAAAAAATGGCCAGGTTGAGCCCAAAATTCGCTGAGCGCAATACGTCTCTCAAG 471

Qy 424 GACCTGACCGAGAACCCGAGCCACATGGCGCGCTATGGAGAGATTTATGAGGATCGC 483

Db 472 GACCTGACCGA-AANCCNANCCACCTGGCGGCTATGAAAAAATTTTNGANGATCCC 530

Qy 484 AAGGCCAGGTCCATTGGTGTCTCCAACTGGACCATTCGCGACCTTGAGAAAGATGCCAAG 543

Db 531 AAGGCCAGGTCCATTGGTGTTCCTCAATTTGGACCATTCGCGACCTTGAGAAAGATGCCAAG 590

Qy 544 TTGCGCAAGTTCATGCTCAGCCCAACAGATTCAGAGATTCACCCCTTCTGCCCCAACGAG 603

Db 591 TTGCGCAAGTTCATGCTCAGCCCAACAGATTCAGAGATTCACCCCTTCTGCCCCAACGAG 650

Qy 604 GAGCTGGTGCAGTACTGCTTCTCCAGAACATTTATGCGCGTGGCGCTA 650

Db 651 GAGCTGGTGCAGTACTGCTTCTCCAGAACATTTATGCGCGTGGCGCTA 697

RESULT 6

AAF14590

ID AAF14590 standard; cDNA; 637 BP.

XX AC AAF14590;

XX DT 13-MAR-2001 (first entry)

XX Aspergillus oryzae EST SEQ ID NO:7113.

XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.

XX Aspergillus oryzae.

XX WO200056762-A2.

PN

Db	301	TCCTGTGAAGCTTTGGACTTGAATGATGTTCTCGTTACTGGCCATTGCT	360
Qy	361	GCGAAGAATGGCCAGGTGAGCCCAAGATTGCGCCTGACGCGAAATACGTATTCTC	420
Db	361	GCGAAGAATGGCCAGGTGAGCCCAAGATTGCGCCTGAGCGAAATACGTATTCTC	420
Qy	421	AAGGACCTCACCAGAAACCCGAGCCACATGGCGCCTATGGAGAGATTATGAGAT	480
Db	421	AAGGACCTGACCGAARACCCGAGCCACATGGCGCGCTATGGAGAGATTATGAGAT	480
Qy	481	CGAAGCCAGGTCCATTGGTGTCTCCAACTGGACCAATGCCGACCTTGAGAAGTGTCC	540
Db	481	CGAAGCCAGGTCCATTGGTGTCTCCAACTGGACCAATGCCGACCTTGAGAAGTGTCC	540
Qy	541	AAGTTGCGCAAGGTATGCTCATGCCAACACAGATCGAGATTCACCCCTTCCTGCCCAAC	600
Db	541	AAGTTGCGCAAGGTATGCTCATGCCAACACAGATCGAGATTCACCCCTTCCTGCCCAAC	600
Qy	601	GAGGAGCTGGTCCAGTACTGCTTCTCCAAAGAACATTATGCCCGTGGCTACTCTCCTCTG	660
Db	601	GAGGAGCTGGTCCAGTACTGCTTCTCCAAAGAACATTATGCCCGTGGCTACTCTCCTCTG	660
Qy	661	GGCTCGCAGAACACAGTTCCACACACCGGTGAGCGGGTCAGCGAGAACAGACTCTGTAAC	720
Db	661	GGCTCGCAGAACAGTTCCACACACCGGTGAGCGGGTCAGCGAGAACAGACTCTGTAAC	720
Qy	721	GAGATCGCGAAGAGGGGGCAACACCTTCTCAGGTTCTTATTCGTGGGTCTGGCC	780
Db	721	GAGATCGCGAAGAGGGGGCAACACCTTCTCAGGTTCTTATTCGTGGGTCTGGCC	780
Qy	781	CGTGGCTACGTCTGTTCTCCCAAGAGTCCAAACCCCAAGCGCATTGAGTCCCACTTCAAG	840
Db	781	CGTGGCTACGTCTGTTCTCCCAAGAGTCCAAACCCCAAGCGCATTGAGTCCCACTTCAAG	840
Qy	841	AGCATTTAGCTCTCCGATCCGACTTTTGAAGCCATCAATGCCGTTGCCAAGGTCGTCTAC	900
Db	841	AGCATTTAGCTCTCCGATCCGACTTTTGAAGCCATCAATGCCGTTGCCAAGGTCGTCTAC	900
Qy	901	TTCCGTTTCGTCAACATGAAGGATCTTTCCGATATGATGTGCGCCCGAGGAGACCGCC	960
Db	901	TTCCGTTTCGTCAACATGAAGGATCTTTCCGATATGATGTGCGCCCGAGGAGACCGCC	960
Qy	961	AAGAACCTGCTCGGTGA	978
Db	961	AAGAACCTGCTCGGTGA	978

RESULT 4	
ABL59397	
ID	ABL59397 standard; DNA; 996 BP.
XX	
XX	
AC	ABL59397;
XX	
DT	
XX	22-OCT-2002 (first entry)
XX	
DE	DNA sequence of a protein producing (S)-4-bromo-3-hydroxybutanoate.
XX	
XX	(S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
KW	agrochemical; 4-cyano-3-hydroxybutanoic acid; gene; ss;
KW	

PR	07-DEC-2000; 2000JP-00372704.	
PR	15-JAN-2001; 2001JP-00006114.	
PR	02-FEB-2001; 2001JP-00026594.	
PR	11-JUN-2001; 2001JP-00175175.	
XX	(SUMO) SUMITOMO CHEM CO LTD.	
PA		
XX		
PI	Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;	
XX		
XX	WPI; 2002-550350/59.	
DR	P-PSDB; ABB77965.	
XX		
PT	New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by	
PT	asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals	
PT	and agrochemicals.	
XX		
PS	Example 1; Page 43-45; 56pp; English.	
XX		
CC	The present sequence encodes a Penicillium citrinum protein which is	
CC	capable of producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically	
CC	reducing 4-bromo-3-oxobutanoate. The protein and polynucleotides are	
CC	useful for producing optically active (S)-4-bromo-3-hydroxybutanoate,	
CC	which is useful as an intermediate in the production of pharmaceuticals	
CC	and agrochemicals. The (S)-4-bromo-3-hydroxybutanoate produced can also	
CC	be used to produce 4-cyano-3-hydroxybutanoic acid	
XX		
SQ	Sequence 996 BP; 226 A; 285 C; 262 G; 223 T; 0 U; 0 Other;	
Query Match 100.0%; Score 978; DB 6; Length 996;		
Best Local Similarity 100.0%; Pred. No. 1.7e-252;		
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 ATGTCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGCGTCGGCTTT	60
DB	1 ATGTCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGCGTCGGCTTT	60
QY	61 GGTACTTCGCTAGTGAAGTTTCCAAAGGCGAGACTTATCTGTCACCATGCGCCTG	120
DB	61 GGTACTTCGCTAGTGAAGTTTCCAAAGGCGAGACTTATCTGTCACCATGCGCCTG	120
QY	121 AAGACGGTTACCGTCACTTCGACTGTGCTCTGCTACTCTGAACGAGGGTGAGGTGGT	180
DB	121 AAGACGGTTACCGTCACTTCGACTGTGCTCTGCTACTCTGAACGAGGGTGAGGTGGT	180
QY	181 GAGGGTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC	240
DB	181 GAGGGTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC	240
QY	241 TGCACCAAGGTGTGAACCACTCCACGGTATGAGGACGTCCTCTGTTCCATTGACGAC	300
DB	241 TGCACCAAGGTGTGAACCACTCCACGGTATGAGGACGTCCTCTGTTCCATTGACGAC	300
QY	301 TCCTCTGAAGCGTCTTGGACTTGACTAGTTGATATGTTCTCGTTCACCTGGCCCATTTGCT	360
DB	301 TCCTCTGAAGCGTCTTGGACTTGACTAGTTGATATGTTCTCGTTCACCTGGCCCATTTGCT	360
QY	361 GCGGAGAGATATGCGCCAGGGTGAGCCCAAGATTGGCCCTGACGGGAAATACGTCATTCTC	420
DB	361 GCGGAGAGATATGCGCCAGGGTGAGCCCAAGATTGGCCCTGACGGGAAATACGTCATTCTC	420
QY	421 AAGGACCTTGACCGAGAACCCGAGCCACATGCGCGCTATGAGAGAGATTTATGAGGAT	480
DB	421 AAGGACCTTGACCGAGAACCCGAGCCACATGCGCGCTATGAGAGAGATTTATGAGGAT	480
QY	481 CGCAAGGCCAGGTCCATTGGTGTCTCCAACTGGACCAATTGCCGACCTTGAGAGATGTCTC	540
DB	481 CGCAAGGCCAGGTCCATTGGTGTCTCCAACTGGACCAATTGCCGACCTTGAGAGATGTCTC	540
QY	541 AAGTTGCGCAAGGTATCGCTCAGCCCAACCAAGATCGAGATTCAGCCCTTCCTGCCCAAC	600
DB	541 AAGTTGCGCAAGGTATCGCTCAGCCCAACCAAGATCGAGATTCAGCCCTTCCTGCCCAAC	600
QY	601 GAGGAGCTGGTGAGTACTGCTTCTCCAAGAACATTATGCCGTGGCTACTCTCTCTGTG	660

121 AAGACCGGTTACCGTCACTTGGACTGCTGCTGTACTACCTGAACGAGGTTGGT 180
181 GAGGATATCGGTACTTCTTGAAGGAGAACCCCTCGGTGAAGGTGAGGACATCTTCGTC 240
181 GAGGATATCGGTACTTCTTGAAGGAGAACCCCTCGGTGAAGGTGAGGACATCTTCGTC 240
241 TGCACCAAGGTGTGAACCACTCCACCGTTATGAGGACGTCCTCTGTCATTCAGCAG 300
241 TGCACCAAGGTGTGAACCACTCCACCGTTATGAGGACGTCCTCTGTCATTCAGCAG 300
301 TCCCTGAGCGTCTGGACTTGAATGACTGATGATGCTTCTGCTCACTGCGCCATTCGCT 360
301 TCCCTGAGCGTCTGGACTTGAATGACTGATGATGCTTCTGCTCACTGCGCCATTCGCT 360
361 GCCGAGAGAAATGCGGAGGTGAGCCCAAGATTGGCCCTGACGCCAAATACGTCATTC 420
361 GCCGAGAGAAATGCGGAGGTGAGCCCAAGATTGGCCCTGACGCCAAATACGTCATTC 420
421 AAGACCTGACCGAGAACCCCGAGCCGACATGCGCGCTATGCGGAGAGATTTATGAGGAT 480
421 AAGACCTGACCGAGAACCCCGAGCCGACATGCGCGCTATGCGGAGAGATTTATGAGGAT 480
481 CGCAGGCGAGTCTGATGCTTCTCAACTGGACCATTCGCGACCTTGAGAAGATGTC 540
481 CGCAGGCGAGTCTGATGCTTCTCAACTGGACCATTCGCGACCTTGAGAAGATGTC 540
541 AAGTTGCGCAAGGTGATGCTTCTCAACTGGACCATTCGCGACCTTGAGAAGATGTC 600
541 AAGTTGCGCAAGGTGATGCTTCTCAACTGGACCATTCGCGACCTTGAGAAGATGTC 600
601 GAGGAGCTGTCGAGTACTGCTTCTCAAGAACATTTGCGCGTGGCTACTCTCTCTG 660
601 GAGGAGCTGTCGAGTACTGCTTCTCAAGAACATTTGCGCGTGGCTACTCTCTCTG 660
661 GCCTCGCAGAACCGAGTTCCCAACCGGTGAGCGGTGAGCGGAGAACAGACTCTGAAC 720
661 GCCTCGCAGAACCGAGTTCCCAACCGGTGAGCGGTGAGCGGAGAACAGACTCTGAAC 720
721 GAGATCGCGAGAGGCGGCAACACCTTCTGAGTCTTATGCTGGGTCTGCGC 780
721 GAGATCGCGAGAGGCGGCAACACCTTCTGAGTCTTATGCTGGGTCTGCGC 780
781 CGTGCTACGTCGTTCTCCCAAGAGCTCCCAACCGGTGAGCGGTGAGCGGAGAACAG 840
781 CGTGCTACGTCGTTCTCCCAAGAGCTCCCAACCGGTGAGCGGTGAGCGGAGAACAG 840
841 AGCATGAGCTCTCGATGCGGACTTTGAAGCCATCAATGCGGTGCGAGGTCGTCAC 900
841 AGCATGAGCTCTCGATGCGGACTTTGAAGCCATCAATGCGGTGCGAGGTCGTCAC 900
901 TTCGTTTCGTCACATCAAGGATACTTTCGATATGATGCTGCGCGGAGGAGACCGCC 960
901 TTCGTTTCGTCACATCAAGGATACTTTCGATATGATGCTGCGCGGAGGAGACCGCC 960
961 AAGAACCTGTCGCGTGA 978
961 AAGAACCTGTCGCGTGA 978

RESULT 3
ADE39630
ID ADE39630 standard; DNA; 978 BP.
XX ADE39630;
XX ADE39630;
DT 29-JAN-2004 (first entry)
XX Penicillium reducing enzyme protein encoding DNA, SEQ ID No 4.
DE Penicillium reducing enzyme protein encoding DNA, SEQ ID No 4.
XX optically active 2-hydroxycycloalkane carboxylic acid ester;
KW 2-oxocycloalkane carboxylic acid ester; enzyme; Penicillium; reduction;
KW gene; ds.

XX Penicillium citrinum.
OS
XX Key Location/Qualifiers
PH 1. .978
FT /*tag= a
FT /product= "Penicillium reducing enzyme protein"
XX
XX EP1323827-A2.
XX 02-JUL-2003.
XX
XX 20-DEC-2002; 2002EP-00258814.
XX
XX 27-DEC-2001; 2001JP-00395884.
XX 27-DEC-2001; 2001JP-00395885.
XX 10-APR-2002; 2002JP-00107648.
XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX
XX Asako H, Wakita R, Itoh N;
PI WPI; 2003-723302/69.
XX P-PSDB; ADE39629.
XX
XX Producing optically active 2-hydroxycycloalkane carboxylic acid ester by
XX reacting 2-oxocycloalkane carboxylic acid ester with transformant having
XX ability to asymmetrically reduce the ester to optically active ester.
XX
XX Claim 9; SEQ ID NO 4; 45pp; English.
XX
XX The invention relates to a novel method for producing an optically active
XX 2-hydroxycycloalkane carboxylic acid ester. The novel method involves:
XX allowing 2-oxocycloalkane carboxylic acid ester to react with a
XX transformant, or a dead cell or their extract, artificially provided with
XX the ability to asymmetrically reduce a 2-oxocycloalkane carboxylic acid
XX ester to the optically active 2-hydroxycycloalkane carboxylic acid; and an
XX ability to regenerate a coenzyme on which an enzyme having the above
XX ability depends; and collecting the resulting optically active 2-
XX hydroxycycloalkane carboxylic acid. The optically active 2-
XX hydroxycycloalkane carboxylic acid is useful as an intermediate for the
XX production of bioactive substances. This polynucleotide sequence
XX represents the DNA encoding a Penicillium enzyme protein used to reduce 2
XX -oxocycloalkane carboxylic acid ester to the optically active 2-
XX hydroxycycloalkane carboxylic acid as part of the method of the invention.
XX
XX Sequence 978 BP; 218 A; 282 C; 261 G; 217 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 978; DB 9; Length 978;
Best Local Similarity 100.0%; Pred. No. 1.7e-252; Indels 0; Gaps 0;
Matches 978; Conservative 0; Mismatches 0;
QY 1 ATGCTCTAACGGAAGACTTTTCACATTGAGCAACGGCGTCAAGATTCTCGGCTCGGCTTT 60
Db 1 ATGCTCTAACGGAAGACTTTTCACATTGAGCAACGGCGTCAAGATTCTCGGCTCGGCTTT 60
QY 61 GGTACCTTCGCTAGTGAAGGTTCCAAAGGCGAGACCTATCTGCTGTCACCACTGCTG 120
Db 61 GGTACCTTCGCTAGTGAAGGTTCCAAAGGCGAGACCTATCTGCTGTCACCACTGCTG 120
QY 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGCTACTACCTCAAGAGGAGGTGAGGTTGCT 180
Db 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGCTACTACCTCAAGAGGAGGTGAGGTTGCT 180
QY 181 GAGGATATCGGTACTTCTTGAAGGAGAACCCCTCGGTGAAGCGGTGAGGACATCTTCGTC 240
Db 181 GAGGATATCGGTACTTCTTGAAGGAGAACCCCTCGGTGAAGCGGTGAGGACATCTTCGTC 240
QY 241 TGCACCAAGGTGTGAACCACTCCACCGTTATGAGGACGTCCTCTGTCATTCAGCAG 300
Db 241 TGCACCAAGGTGTGAACCACTCCACCGTTATGAGGACGTCCTCTGTCATTCAGCAG 300
QY 301 TCCTGGAAGCGTCTTGGACTTGACTGATGATGCTTCTGCTCACTGCGCCATTCGCT 360

CC useful for producing optically active (S)-4-bromo-3-hydroxybutanoate,
CC which is useful as an intermediate in the production of pharmaceuticals
CC and agrochemicals. The (S)-4-bromo-3-hydroxybutanoate produced can also
XX be used to produce 4-cyano-3-hydroxybutanoic acid
SQ Sequence 978 BP; 218 A; 282 C; 261 G; 217 T; 0 U; 0 Other;

Query Match 100.0%; Score 978; DB 6; Length 978;
Best Local Similarity 100.0%; Pred. No. 1.7e-252;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTAACGGAAGACTTTCACATTGAGCAACGGCTCAAGATTCTCGCGCTGGCTTT 60
DB 1 ATGCTAACGGAAGACTTTCACATTGAGCAACGGCTCAAGATTCTCGCGCTGGCTTT 60
QY 61 GGTACCTTCGTAGTGAAGTTCACAGGGGAGACCTTATCTGTCTGCACCTTGCCTG 120
DB 61 GGTACCTTCGTAGTGAAGTTCACAGGGGAGACCTTATCTGTCTGCACCTTGCCTG 120
QY 121 AAGACCGGTTACCGTCACTTGGACTGTCCCTGTACTACCTGAACGAGGGTGAGTTGGT 180
DB 121 AAGACCGGTTACCGTCACTTGGACTGTCCCTGTACTACCTGAACGAGGGTGAGTTGGT 180
QY 181 GAGGTATCCGTGACTTCTTCAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
DB 181 GAGGTATCCGTGACTTCTTCAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
QY 241 TGCACCAAGGTGTGAACCACTTCAAGGTTATGAGGACGTCCTCTGTCTCATTTGAGGAC 300
DB 241 TGCACCAAGGTGTGAACCACTTCAAGGTTATGAGGACGTCCTCTGTCTCATTTGAGGAC 300
QY 301 TCCCTGAAGCGCTTGGACTTGAATGATGATGTTCTCGTTTCACTGGCCCATTTGCT 360
DB 301 TCCCTGAAGCGCTTGGACTTGAATGATGATGTTCTCGTTTCACTGGCCCATTTGCT 360
QY 361 GCCGAGAGAAATGCGCAGGGTGAAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATCTC 420
DB 361 GCCGAGAGAAATGCGCAGGGTGAAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATCTC 420
QY 421 AAGGACCTGACCGAGAACCCCGAGCCACATCGGCGCTATGAGAGAGATTTAGGAT 480
DB 421 AAGGACCTGACCGAGAACCCCGAGCCACATCGGCGCTATGAGAGAGATTTAGGAT 480
QY 481 CGCAAGCCAGGTCCATTGGTGTCTTCAACTTGGACCAATTGCCGCTTGAAGAGATGCC 540
DB 481 CGCAAGCCAGGTCCATTGGTGTCTTCAACTTGGACCAATTGCCGCTTGAAGAGATGCC 540
QY 541 AAGTTCGCCAGGTTCATGCTCAGCCCAACAGATCGAGATTCAACCCCTTCTGCCCAAC 600
DB 541 AAGTTCGCCAGGTTCATGCTCAGCCCAACAGATCGAGATTCAACCCCTTCTGCCCAAC 600
QY 601 GAGGAGCTGTGCACTGTCTTCCAAAGAACATTATGCCGCTGGCTTCTCTCTCTG 660
DB 601 GAGGAGCTGTGCACTGTCTTCCAAAGAACATTATGCCGCTGGCTTCTCTCTCTG 660
QY 661 GCCTCGAGAACAGGTTCACACACCGGTGAGCGGTGACGAGAACAGACTCTGAAC 720
DB 661 GCCTCGAGAACAGGTTCACACACCGGTGAGCGGTGACGAGAACAGACTCTGAAC 720
QY 721 GAGATCGCCGAGAGGGCGGCAACACCTTCTCAGGTTCCTTATTGCTGGGTCTGCGC 780
DB 721 GAGATCGCCGAGAGGGCGGCAACACCTTCTCAGGTTCCTTATTGCTGGGTCTGCGC 780
QY 781 CGTGGTACGTCTTCCCAAGAGTTCACACCCCAAGCGCATTTAGTCCAACTTCAAG 840
DB 781 CGTGGTACGTCTTCCCAAGAGTTCACACCCCAAGCGCATTTAGTCCAACTTCAAG 840
QY 841 AGCATGAGCTCTCGATCGCGACTTTGAAGCCATCAATGCCGTTGCAAGGGTCTGTCAC 900
DB 841 AGCATGAGCTCTCGATCGCGACTTTGAAGCCATCAATGCCGTTGCAAGGGTCTGTCAC 900
QY 901 TTCGGTTCTGTCACATGAAGGATCTTTGGATATGATGTCTGGCCCGGAGGACCGCC 960
DB 901 TTCGGTTCTGTCACATGAAGGATCTTTGGATATGATGTCTGGCCCGGAGGACCGCC 960

DB 901 TTCGGTTCTGTCACATGAAGGATCTTTGGATATGATGTCTGGCCCGGAGGACCGCC 960
QY 961 AAGAACTCTCTGCGTGA 978
DB 961 AAGAACTCTCTGCGTGA 978

RESULT 2
ABL59398
ID ABL59398 standard; DNA; 978 BP.

XX ABL59398;
AC
XX
XX
DT 22-OCT-2002 (first entry)

XX DNA sequence of a protein producing (S)-4-bromo-3-hydroxybutanoate.

XX (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
KW agrochemical; 4-cyano-3-hydroxybutanoic acid; gene; ss.

XX Penicillium citrinum.

XX Key Location/Qualifiers
FH 1. 978
FT /*tag= a
FT
XX

PN EP1213354-A2.

XX 12-JUN-2002.

XX 07-DEC-2001; 2001EP-00310251.

XX 07-DEC-2000; 2000JP-00372704.

PR 15-JAN-2001; 2001JP-00006144.

PR 02-FEB-2001; 2001JP-00026594.

PR 11-JUN-2001; 2001JP-00175175.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;

XX WPI; 2002-550350/59.

DR P-PSDB; ABB77965.

XX New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by
PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals
PT and agrochemicals.

PS Example 5; Page 45-47; 56pp; English.

XX The present sequence encodes a Penicillium citrinum protein which is
CC capable of producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically
CC reducing 4-bromo-3-oxobutanoate. The protein and polynucleotides are
CC useful for producing optically active (S)-4-bromo-3-hydroxybutanoate,
CC which is useful as an intermediate in the production of pharmaceuticals
CC and agrochemicals. The (S)-4-bromo-3-hydroxybutanoate produced can also
CC be used to produce 4-cyano-3-hydroxybutanoic acid

XX Sequence 978 BP; 218 A; 282 C; 261 G; 217 T; 0 U; 0 Other;

Query Match 100.0%; Score 978; DB 6; Length 978;

Best Local Similarity 100.0%; Pred. No. 1.7e-252;

Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTAACGGAAGACTTTCACATTGAGCAACGGCTCAAGATTCTCGCGCTGGCTTT 60

DB 1 ATGCTAACGGAAGACTTTCACATTGAGCAACGGCTCAAGATTCTCGCGCTGGCTTT 60

QY 61 GGTACCTTCGTAGTGAAGTTCACAGGGGAGACCTTATCTGTCTGCACCTTGCCTG 120

DB 61 GGTACCTTCGTAGTGAAGTTCACAGGGGAGACCTTATCTGTCTGCACCTTGCCTG 120

QY 121 AAGACCGGTTACCGTCACTTGGACTGTGCTGGTACTACCTGAACGAGGGTGAGTTGGT 180

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2004, 09:47:51 ; Search time 446 Seconds
(without alignments)
9315.555 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978
Sequence: 1 atgtctaacggaaagacttt.....ccaagaacctgtctgctga 978

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1930s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	978	100.0	978	6	ABL59376
2	978	100.0	978	6	ABL59398
3	978	100.0	978	9	Abg39630 Penicilli
4	978	100.0	996	6	ABL59397
5	588.6	60.2	697	6	ABL59384
6	420.2	43.0	637	3	AAFL4590
7	417	42.6	417	6	ABL59391
8	415.4	42.5	587	7	ABZ53802
9	412	42.1	743	6	ABL59388
10	265	27.1	331	6	ABL59387
11	199.6	20.4	972	2	AAFL2160
12	199.6	20.4	1055	2	AAFL2159
13	199.6	20.4	1055	2	AAFL2159
14	178.8	18.3	978	4	ABL10491
15	176.4	18.0	403	7	ABZ56044
16	144.2	14.7	1121	6	ABQ82227
17	142.6	14.6	1077	6	ABQ82226
18	135.4	13.8	656	7	ABZ53102
19	131.8	13.5	628	7	ABZ53102
20	129.4	13.2	1385	3	AAFL07918
21	127.2	13.0	963	4	ABL16915
22	120.8	12.4	1394	2	AAQ05879
23	119.2	12.2	1367	6	ABK09795

24	119.2	12.2	1367	7	ABZ71966
25	119.2	12.2	1368	6	ABL58970
26	119	12.2	951	6	ABL01106
27	118	12.1	1292	4	ABL05161
28	117	12.0	1076	4	ABL16913
29	116.2	11.9	1059	7	ADA70389
30	116	11.9	1371	2	AAQ14946
31	115.2	11.8	1380	5	AAZ58973
32	113.2	11.6	942	5	AAH74584
33	113	11.6	1337	8	AAV58414
34	113	11.6	1337	8	ACF25340
35	113	11.6	1339	7	ABT41790
36	112	11.5	1163	3	AAZ50286
37	112	11.5	1163	3	AAZ45937
38	111.8	11.4	1913	3	AAZ77653
39	111	11.3	972	8	AAZ57091
40	111	11.3	972	8	AAZ57086
41	111	11.3	972	8	AAZ57088
42	111	11.3	972	9	AAZ42952
43	111	11.3	1290	8	AAZ57083
44	109.4	11.2	1064	8	AAZ57087
45	109.4	11.2	1165	3	AAZ33951

ALIGNMENTS

RESULT 1

ABL59376
ID ABL59376 standard; DNA; 978 BP.

XX
AC ABL59376;

XX
DT 22-OCT-2002 (first entry)

XX
DE DNA sequence of a protein producing (S)-4-bromo-3-hydroxybutanoate.

XX
KW (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;

XX
KW agrochemical; 4-cyano-3-hydroxybutanoic acid; gene; ss.

XX
OS Penicillium citrinum.

XX
FH Key Location/Qualifiers

FT CDS 1..978

FT /*tag= a

XX
PN EP1213354-A2.

XX
PD 12-JUN-2002.

XX
PF 07-DEC-2001; 2001EP-00310251.

XX
PR 07-DEC-2000; 2000JP-00372704.

PR 15-JAN-2001; 2001JP-00006144.

PR 02-FEB-2001; 2001JP-00026594.

PR 11-JUN-2001; 2001JP-00175175.

XX
PA (SUMO) SUMITOMO CHEM CO LTD.

XX
PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;

XX
DR WPI; 2002-550350/59.

DR P-PSDB; ABB77965.

XX
PT New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by

PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals

XX
PT and agrochemicals.

XX
PS Claim 1; Page 31-33; 56pp; English.

XX
CC The present sequence encodes a Penicillium citrinum protein which is

CC capable of producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically

CC reducing 4-bromo-3-oxobutanoate. The protein and polynucleotides are

362	QY	CGAGAGAAATGGCCAGGGTGAGCCCAAGATTGGCCCTGAGCGCAAAATACGTCTATTCTCA	421
245	Db	CGGAGAGAACCGACGACTATGAAGTCAAGATCGGAGATGATGAAAGTAGTACATCATCAACA	304
422	QY	AGGACCTCAGCCGAGAAACCCGAGGCCCATATGGCGCGCTATGGAGAGAGATTTATGAGGATC	481
305	Db	AAGAGTCAACGGCCAACTTGGAACTATTGGGGTAGATTCGAGGCTCTTAACAAGCTG	364
482	QY	GCAAGGCCAGGTCANATTGGTGTCTCCAACTGGACCAATTGCCGACCTTGAGAGAGATGTCCA	541
365	Db	GCAAGGCAAAAGCCATCGGAGTATCCAAATTTTACCATCTCAAAACCTTGGAGGCTCTTCTCA	424
542	QY	AGTTTCGCCAAAGTCATGCTCACGCCCAACAGATTCGAGATTACCCCTTCTCTGCCCAACG	601
425	Db	AGTAGCGAGATGTTCTCTCCGCCATCAACCAAGTCGAGATTATCCGCTCTGGCCCTACA	484
602	QY	AGGAGCTGGTGAGTACGTCTTCCAAAGAACATTAATGCCCCGTGGCCTACTCTCTCTGG	661
485	Db	CTAAGTTGATCAACTACTCTGTTTTTCAAAGAACATCTCTTCCGTGGCATATTCCGCACCTG	544
662	QY	GCTCCGACAAACAGGTTCCCAACACCGGTGAGCGGGTCAAGCGAGAACAGACTCTGAAAG	721
545	Db	GATCCAGAGCCAGTGCCCAAGCAGACGCAAAACCGTGATCCAGAACTCTGAACTTAICT	604
722	QY	AGATCCCGCAGAGGGCGCAACACCCTTGCTCAGGTTCTTAATGGCTGGGTTGCGCC	781
605	Db	TAATTTCGGGAAAAGAGGGCGTTAGCATAGTCAAAATATTGATGCTTGTGTGTATCAAAC	664
782	QY	G-TGGCTACGTGTCTCCCAACAGAGTCCAAACCCCAAGGCAATTCAGTCCAACTTCAAG	840
665	Db	GAAAGATTAACCTAATCTTCCCATGAGTCAATGAGGGGAGAAATCAAAACCAATGGNACA	724
841	QY	AGCATTTAGCTCTCCGATGCCGACTTTGAAGCCCATCAATGCCGTTGCCAAG	891
725	Db	CTTGTGCAATTGACTGCAAGAGGGTTCCGAAAAAATACCCCAAGTCCGAGG	775

Search completed: May 29, 2004, 12:17:05
Job time : 3881 secs

```
Qy 620 GCTTCTCAAGAACATTATGCGCTGCTACTCTCC 656
|||
Db 684 GCCTTTCGAAGATATTGTACCACTTGCATATCCCC 720

RESULT 14
CNS01BYU 660 bp mRNA linear PLN 02-SEP-1999
LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
DEFINITION nitrogen deprivation.
ACCESSION AL114862
VERSION AL114862.1 GI:5829481
KEYWORDS cDNA library; nitrogen deprivation.
SOURCE Botryotinia fuckeliana
ORGANISM Botryotinia fuckeliana
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
Helotiales; Sclerotiniaceae; Botryotinia.
1 (bases 1 to 660)
REFERENCE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
AUTHORS Direct Submission
TITLE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
JOURNAL 78026 Versailles, France
REFERENCE 2 (bases 1 to 660)
Genoscope.
Direct Submission
TITLE Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
JOURNAL CP 5706 91057 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT The cDNA library to be analyzed within the framework of this
project was created using a Botrytis cinerea strain which was grown
under conditions of nitrogen deprivation, which is the normal
situation for B. cinerea during its development on its host plant.
The library was produced in an oriented direction, in the pBSII
vector.
FEATURES Location/Qualifiers
source 1..660
/organism="Botryotinia fuckeliana"
/mol_type="mRNA"
/strain="T4"
/db_xref="taxon:40559"
/note="Genoscope sequence ID : W71D061"
ORIGIN
Query Match 23.6%; Score 230.6; DB 8; Length 660;
Best Local Similarity 62.3%; Pred. No. 4.le-39;
Matches 359; Conservative 0; Mismatches 217; Indels 0; Gaps 0;
20 TCACATTGAGCAACGGGTCAAGATTCTGCGTCGGCTTGTGACCTTCGCTAGTGAAG 79
|||
Db 85 TCACAGCTGAACAATGGAGTTCAGATGCCAGCTTTTGGTTTCGGCACATTTGCCAGTGAGG 144
|||
Qy 80 GTTCCAGGGCGAGACCTATCTGTCACCTGCCCTGAAGACCGGTTACCGTCACT 139
|||
Db 145 GTACGGTAGGAAGACACACACAGCCGGCTGTAGCAGTTTGAACGGANGATACAGACATT 204
|||
Qy 140 TGGACTGTGCTGTACTACTGAACAGGGGTGAGGTGTTGGTGGGATCCGTCACCTCC 199
|||
Db 205 TGGACTGTGCTGTGTTCTTACCAAAATGAGAACGAAGTCGGAACCGCTTAAAGAGTTCC 264
|||
Qy 200 TGAAGGAGAACCCCTCGTGAAGCGTGAAGGACATCTTCGTCGACCAAGGTGGAAC 259
|||
Db 265 TTGCTGCCAACCCCAAGTGTGAAGAGATCCGACATTTTCATTTGTCACAAAGGTCTGGATC 324
|||
Qy 260 ACCTCCACCGTTATGAGGACGCTCTCTGGTCCATTGACGACTCCCTGAAGCGTCTTGGAC 319
|||
Db 325 ACCTCCACGACCTGAAGATGTGAGTGGAGCTTGAAGAACTCGTTAGAAAGCTTCAAA 384
|||
Qy 320 TTGACTAGTTGATATGTTCTCTCTGCTTCACTGGCCCATTTGTCGCGAAGATGGCCAGG 379
|||
Db 385 CCGCTTACATTGATGTCATTCTTGTGTACACTGGCCAAATGCTGCANAGAAGATGAGACA 444
|||
Qy 380 GTGAGCCCAAGATTGGCCCTGACCGGCAAAATACGTCATTCTCAAGGACCTGACCGAGAACC 439
|||
```

```
Db 445 GATCAGTCAAGATCGGTGCTGATGGAAGTAGTCTCATCAAGAAAGATTGACAGAAAACC 504
|||
Qy 440 CCGAGCCACATGGCGCGCTATGGAGAAGATTATTAGAGATCGCAAGCCAGGTCCATTG 499
|||
Db 505 CAGAGCCACATGGAGAGCATGGANAAATTTATCAAGAGGGACTCCGGAAGGTAATCG 564
|||
Qy 500 GTGTCTCCAATGGACCAATTGCCGACCTTGAGAGATCTCAAAGTTCCCAAGGTCTATGC 559
|||
Db 565 GTGTCTCCAATGGACCGAGGAAGGAATTGAGCAGCTTTTGTCAATCGTCTGAGTTAAGC 624
|||
Qy 560 CTCACGCCAACCCAGATCGAGATTCACCCCTTCCTGC 595
|||
Db 625 CAACAATTACCAATCGAGATCCATCTCTTCCTTC 660
|||

RESULT 15
CNS01BQ2 780 bp mRNA linear PLN 02-SEP-1999
LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
DEFINITION nitrogen deprivation.
ACCESSION AL114546
VERSION AL114546.1 GI:5829165
KEYWORDS cDNA library; nitrogen deprivation.
SOURCE Botryotinia fuckeliana
ORGANISM Botryotinia fuckeliana
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomyces;
Helotiales; Sclerotiniaceae; Botryotinia.
1 (bases 1 to 780)
REFERENCE Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
AUTHORS Direct Submission
TITLE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
JOURNAL 78026 Versailles, France
REFERENCE 2 (bases 1 to 780)
Genoscope.
Direct Submission
TITLE Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
JOURNAL CP 5706 91057 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT The cDNA library to be analyzed within the framework of this
project was created using a Botrytis cinerea strain which was grown
under conditions of nitrogen deprivation, which is the normal
situation for B. cinerea during its development on its host plant.
The library was produced in an oriented direction, in the pBSII
vector.
FEATURES Location/Qualifiers
source 1..780
/organism="Botryotinia fuckeliana"
/mol_type="mRNA"
/strain="T4"
/db_xref="taxon:40559"
/note="Genoscope sequence ID : W35C041"
ORIGIN
Query Match 23.1%; Score 226.2; DB 8; Length 780;
Best Local Similarity 57.6%; Pred. No. 3.5e-38;
Matches 444; Conservative 0; Mismatches 323; Indels 4; Gaps 2;
125 CCGGTTACCGTCACTTGGACTGTGCTGCTACTACTCTCAACGAGGCTGAGGTGTTGTTGAGG 184
|||
Db 5 CAGGTTATCGCCATCTAGATTCCGCTTGGTATTACAAGATGAAGAGAGGTTGGTCTG 64
|||
Qy 185 GTATCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCTGAGGACATCTTCGTTCTCA 244
|||
Db 65 GCTTCGGGAAATCTTGTCCAGCAACCCCAAGGTCAAGCGTTCCGATATCTTCATCACCA 124
|||
Qy 245 CCAAGGTGGGAACCACTCCACCGTT--ATGAGGAGTCTCTCTGGTCCATTTGACGACT 301
|||
Db 125 CGAAGTGTGGCGCTCATCTATCGGTTCTCTCGAAGAGCTCGAGTGGAGTCTGATTACA 184
|||
Qy 302 CCGTGAAGCGTCTTGACCTTGACTAGTTGATATGTTCTCTCGTTCACTGGCCCATTTGCTG 361
|||
Db 185 GTTAGAGAAACTAGGGGTGATATGTGGATTGTGGATTCTTTCTTGATGATTTGGCCATTCGCG 244
|||
```

COMMENT
The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES
source
1. 720
/organism="Botryotinia fuckeliana"
/mol_type="mRNA"
/strain="T4"
/db_xref="taxon:40559"
/note="Genoscope sequence ID : W26B011"

ORIGIN
Query Match 25.3%; Score 247; DB 8; Length 720;
Best Local Similarity 62.5%; Pred. No. 1.2e-42;
Matches 402; Conservative 0; Mismatches 240; Indels 1; Gaps 1;
20 TCACATTGAGCAACGGCTCAAGATTCTCGCTGGCTGGCTTTGGTACCTTCGCTAGTGAAG 79
Db TCAAGCTGAACAATGGAGTTCAGATGCCAGCTTTGGTTTCGGACATTTGCCAGTGAGG 129
Qy 80 GTTCCAAAGGGGAGACCTATCTCTGTACACCTGCTGCAAGACCGGTACCGTCACT 139
Db GTACGGTAGGAGACACACAGAGCCCTGTAGCAGCTTTGAACGAGGATACAGACATT 189
Qy 140 TGGACTGTGCTGTACTACTGTAACGAGGGTGAGGTGTGGTGGGTATCCGTGACTTC 199
Db TGGACTGTGCTGTGTTCTACCAATAGAGAACGAAGTCGGAACGCGGTTAAGGAGTTC 249
Qy 200 TGAAGGAGAACCCCTCGGTGAAGGGTGAGGACATCTTCGTCTGACCAAGCTGTGAACC 259
Db TTGCTGCCAACCAAGTGTGAGAGATCCGACATTTTCATGTTCACAAAGTCTGGAATC 309
Qy 260 ACCTCCACCGTATGAGGAGCTCTCTGTGTCATGACAGCTCCCTGAAGCGTCTTGGAC 319
Db ACCTCCACGACCTGAAGATGTCGAGTGGAGCTTGAAGAACTCGTTAGAAAAGCTTCAA 369
Qy 320 TTGACTAGTTGATATGTTCTCGTCTGCTACCTGGCCATTTGTCGCGAGAAATGGCCAG 379
Db CCCCTTACATGATGATCTTTGTTACACTGGCCAAATGCTGAGAGAAATGAGGACAG 429
Qy 380 GTGAGCCCAAGATTGGCCCTGACGGCAATACGTCATTCTCAAGGACCTGACCGAGAAC 439
Db ATCAG-TCAGGATCGGTCTGATGGAAGTAGGTTCATCAGGAAAGATTTCACAGAAAAC 488
Qy 440 CCGAGCCACATGCGGCGCTATGAGAGAGATTATGAGGATCGCAAGGCCAGTCCATTG 499
Db CAGAGCCCAACATGAGAGCAATGAGAAAGTTATACAGAGGAGGACTCGCAAGGCTATCG 548
Qy 500 GTGCTCCAACTGACCATTCGCCACCTTCGAGAGATGTCCAAGTTCGCCCAAGGCTATGC 559
Db GTGCTCCAACTGACCGAGAGGAAATGAGAGCTTTTGTCTATTCGCTGAGGTTAAGC 608
Qy 560 CTCAGGCCAACCGATCGAGATTCAACCCCTTCCTGCCCCAACGAGGAGCTGGTGCAGTACT 619
Db CAACAATTAACCAATCGAGATCATCTCTTCTTCCACAAACAAAGCTCATCGACTACT 668
Qy 620 GCTTCTCCAGAACATTATGCGGTGGCGCTACTCTCTCTCTGGG 662
Db GCCTTTTGAAGATATTGTACCAAGTTCGATATTCCTCCATTTGGG 711

RESULT 13
CNS019U1 Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.
LOCUS AL112097
DEFINITION AL112097.1 GI:5826716
ACCESSION
VERSION
KEYWORDS cDNA library; nitrogen deprivation.
SOURCE Botryotinia fuckeliana

ORGANISM
Botryotinia fuckeliana
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
Helotiales; Sclerotiniaceae; Botryotinia.
REFERENCE
1 (bases 1 to 720)
AUTHORS Bittou, F., Levis, C., Fortini, D., Pradier, J.M. and Brygoo, Y.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequençage : CP 5706 91057 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr

COMMENT
The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES
source
1. 720
/organism="Botryotinia fuckeliana"
/mol_type="mRNA"
/strain="T4"
/db_xref="taxon:40559"
/note="Genoscope sequence ID : W11E111"

ORIGIN
Query Match 23.8%; Score 233.2; DB 8; Length 720;
Best Local Similarity 61.4%; Pred. No. 1.1e-39;
Matches 391; Conservative 0; Mismatches 244; Indels 2; Gaps 1;
Qy 20 TCACATTGAGCAACGGCTCAAGATTCTCGCTGGCTGGCTTTGGTACCTTCGCTAGTGAAG 79
Db TCAAGCTGAACAATGGAGTTCAGATGCCAGCTTTGGTTTCGGACATTTGCCAGTGAGG 145
Qy 80 GTTCCAAAGGGGAGACCTATCTCTGTCAACCTGCTGCAAGACCGGTACCGTCACT 139
Db GTAGCGTAGGAGACACACAGGCGCTGTAGCAGCTTTGAACGAGGATACAGACATT 205
Qy 140 TGGACTGTGCTGTACTACTGTAACGAGGGTGAGGTGTGGTGGGTATCCGTGACTTC 199
Db TGGACTGTGCTGTGTTCTACCAATGAGAACGAGTGGAAACGCGGTTAAGGAGTTC 265
Qy 200 TGAAGGAGAACCCCTCGGTGAAGCGTGAAGGACATCTTCGTCTGCAACCAAGGTGTGGAACC 259
Db TTGCTGCCAACCCAGTGTGAAGATCCGACATTTTCATTTGTCAAAAGCTTGAATC 325
Qy 260 ACCTCCACCGTATGAGGAGCTCTCTGTGTCATGAGACTCCCTGAAGCGTCTTGGAC 319
Db ACCTCCACGACCTGAAGATGTCGAGTGGAGCTTGAAGAACTCGTTAGAAAAGCTTCAA 385
Qy 320 TTGACTAGTTGATATGTTCTCGTCTGCTACCTGGCCATTTGTCGCGAGAAATGGCCAGG 379
Db CCCCTTACATGATGATCTTTGTTACACTGGCCAAATGCTGAGAGAAATGAGGAC 445
Qy 380 GTGAGCCCAAGATTGGCCCTGACGGCAATACGTCATTCTCAAGGACCTGACCGAGAAC 439
Db GATCAGTCAAGATCGGTCTGATGGAAGTACGTCATCAAGAAAGATTTCAGAGAAAACC 505
Qy 440 CCGAGCCCACTGCGCGCTATGAGAGATTTATGAGGATCGCAAGGCCAGGTCATTG 499
Db CAGAGCCCACTGAGAGCAATGGAGAG--TATCAAGAGGAGGACTCGCAAGCTATCG 563
Qy 500 GTGCTCCAACTGAGACCATTCGCCACCTTGAGAAGATGTCGAAGTTCGCCAAGGTCTATGC 559
Db GTGCTCCAACTGAGACCGGAGAGGAAATGAGCACTTTTGTCTATTCGTCAGGTTAAGC 623
Qy 560 CTCAGGCCAACCGATCGAGATTCAACCCCTTCCTGCCAACGAGGAGCTGGTGCAGTACT 619
Db CAACAATTAACCAATTCGAGATCCATCTCTTCTTCCACANCAAAAGCTCATCGACTACT 683

Db 121 AGAAGATGGCCAGGTTGAGCCCAAGATTGGCCCTGACGGCAATATCGTCAATCTCAAGG 180
Qy 425 ACCTGACCGAGAACCCCGAGCCACATGCGCGCTATGGAGAAGATTATGAGATCGCA 484
Db 181 ACCTGACCGAGAACCCCGAGCCACATGCGCGCTATGGAGAAGATTATGAGATCGCA 240
Qy 485 AGGCAGGTTCATTGGTGTCTCCAACTGGACCATTTGCGGACCTTTGAGAAGTGTCCAAGT 544
Db 241 AGGCAGGTTCATTGGTGTCTCCAACTGGACCATTTGCGGACCTTTGAGAAGTGTCCAAGT 300
Qy 545 TCGCAAGGTTCATGCTCAGCCCAACAGATTCAGATTCACCCCTCTCCGCCAACGAGG 604
Db 301 TCGCAAGGTTCATGCTCAGCCCAACAGATTCAGATTCACCCCTCTCCGCCAACGAGG 360
Qy 605 AGCTGGTGCAGTACTGCTTCACCAAGAACATTAATGCGGCGGCTACTCTCCCTGTG 661
Db 361 AGCTGGTGCAGTACTGCTTCACCAAGAACATTAATGCGGCGGCTACTCTCCCTGTG 417

RESULT 10
AX472804
LOCUS AX472804 743 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 19 from Patent EP1213354.
ACCESSION AX472804
VERSION AX472804.1 GI:22207653
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE
1 Asako, H., Matsumura, K., Shimizu, M., Ito, N. and Wakita, R.
TITLE Process for producing optically active 4-halo-3-hydroxybutanoate
JOURNAL Patent: EP 1213354-A 19 12-JUN-2002;
Sumitomo Chemical Company, Limited (JP)
LOCATION/Qualifiers
source 1. 743
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 42.1%; Score 412; DB 6; Length 743;
Best Local Similarity 100.0%; Pred. No. 4.3e-78;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 567 CAACAGATCAGATTACCCCTTCTGCGCCACAGAGAGTGTGCTGAGTACTGCTTCTC 626
Db 1 CAACAGATCAGATTACCCCTTCTGCGCCACAGAGAGTGTGCTGAGTACTGCTTCTC 60
Qy 627 CAAGAACAATTATGCCGTGGCTACTCTCTCTGGCTCGCAGAACCAAGTTCCACCAC 686
Db 61 CAAGAACAATTATGCCGTGGCTACTCTCTCTGGCTCGCAGAACCAAGTTCCACCAC 120
Qy 687 CGGTGAGCGGTGACGAGACAGACTCTCAAGAGATCGCCGAGAGGCGGACAC 746
Db 121 CGGTGAGCGGTGACGAGACAGACTCTCAAGAGATCGCCGAGAGGCGGACAC 180
Qy 747 CTTGCTCAGTCTTATTGCTGGGTCTGCGCGGTGCTGCTGCTTCTCCCGAAGAG 806
Db 181 CTTGCTCAGTCTTATTGCTGGGTCTGCGCGGTGCTGCTGCTTCTCCCGAAGAG 240
Qy 807 CTCACACCCCAAGCGCATTTAGTCAACTCAAGAGATTTAGTCTCCGATCGCGACTT 866
Db 241 CTCACACCCCAAGCGCATTTAGTCAACTCAAGAGATTTAGTCTCCGATCGCGACTT 300
Qy 867 TGAAGCATCAATGCGGTGCAAGGGTCTGCTTCTGCTTCTGCTCAACATGAAGGATAC 926
Db 301 TGAAGCATCAATGCGGTGCAAGGGTCTGCTTCTGCTTCTGCTCAACATGAAGGATAC 360
Qy 927 TTTCGATATGATGTCTGCGCCGAGGAGACCGCCAAAGAACCTGTCTCGGTGA 978
Db 361 TTTCGATATGATGTCTGCGCCGAGGAGACCGCCAAAGAACCTGTCTCGGTGA 412

RESULT 11
AX472803
LOCUS AX472803 331 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 18 from Patent EP1213354.
ACCESSION AX472803
VERSION AX472803.1 GI:22207652
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

ORIGIN

Query Match 27.1%; Score 265; DB 6; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.8e-46;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCTAACGGAAAGACTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGGTCGGCTTT 60
Db 67 ATGTCTAACGGAAAGACTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGGTCGGCTTT 126
Qy 61 GGTACTCTCGTAGTGAAGTTCCAAAGGCGAGACCTACTGCTGCTCACCACCTGCCCTG 120
Db 127 GGTACTCTCGTAGTGAAGTTCCAAAGGCGAGACCTACTGCTGCTCACCACCTGCCCTG 186
Qy 121 AAGACCGGTACCGTCACTTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 187 AAGACCGGTACCGTCACTTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
Qy 181 GAGGTTATCCGTGACTTCTGAGGAGAACCCCTCGGTGAGCGTGAGGACATCTTCGTC 240
Db 247 GAGGTTATCCGTGACTTCTGAGGAGAACCCCTCGGTGAGCGTGAGGACATCTTCGTC 306
Qy 241 TGCACCAAGGTGTGGAACCACTCC 265
Db 307 TGCACCAAGGTGTGGAACCACTCC 331

RESULT 12
CNS01C4M
LOCUS CNS01C4M 720 bp mRNA linear PLN 02-SEP-1999
DEFINITION Botrytis cinerea strain T4 CDNA library under conditions of nitrogen deprivation.
ACCESSION AL115070
VERSION AL115070.1 GI:5829689
KEYWORDS cDNA library; nitrogen deprivation.
SOURCE Botryotinia fuckeliana
ORGANISM Botryotinia fuckeliana
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.

REFERENCE
1 (bases 1 to 720)
AUTHORS Bitton, F., Levis, C., Fortini, D., Pradier, J.M. and Brygoo, Y.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France
REFERENCE 2 (bases 1 to 720)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

```

QY 604 GAGCTGGTGCAGTACTGCTTCTTCCAGAACATTATGCCCCTGGCCTA 650
    |||||
Db 651 GAGCTGGTGCAGTACTGCTTCTTCCAGAACATTATGCCCCTGGCCTA 697
    |||||

RESULT 8
AB014493
LOCUS Gibberella zeae gene for reductase, partial cds. PLN 10-OCT-1998
DEFINITION Gibberella zeae
ACCESSION AB014493
VERSION AB014493.1 GI:3724333
KEYWORDS reductase
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
REFERENCE 1 (sites)
AUTHORS Kimura,M., Matsumoto,G., Shingu,Y., Yoneyama,K. and Yamaguchi,I.
TITLE The mystery of the trichothecene 3-O-acetyltransferase gene.
ANALYSIS Analysis of the region around Trl101 and characterization of its
homologue from Fusarium sporotrichioides
JOURNAL FEBS Lett. 435 (2-3), 163-168 (1998)
MEDLINE 98433864
PUBMED 9762900
REFERENCE 2 (bases 1 to 762)
AUTHORS Kimura,M.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1998) Makoto Kimura, The Institute of Physical
and Chemical Research (RIKEN), Microbial Toxicology Laboratory;
Hirosewa 2-1, Wako, Saitama 351-0198, Japan
(E-mail:mkimura@postman.riken.go.jp, Tel:81-48-467-9518,
Fax:81-48-462-4676)

FEATURES
    source
        1..762
            /organism="Gibberella zeae"
            /mol_type="genomic DNA"
            /strain="F15"
            /db_xref="taxon:5518"
            <1..>762
            /note="putative"
            /codon_start=1
            /product="reductase"
            /protein_id="BA033773.1"
            /db_xref="GI:3724334"
            /translation="PGVGFPTFANEGAKGTYKAVIAALKTYGRHLDCAWFYLNIEGV
            GOAVRDLASDNGVKRDLFICTKWNHLHPREVKSPENSKNLFGLDYLFLVHW
            PIAEKGDYQPKTPGDKYIKDLTENPEPTWAMEEIVASGKARAIGVSNWTIDG
            LQIMAFKVPKPAVNOIEIHPFLPNEDLVKFLNEDLIPAAYPPLSGNQVPTGTIV
            RSNKTLNEVADRSGNTLAQVLIANGURRGLFCFPA"

ORIGIN
    Query Match 48.2%; Score 471; DB 8; Length 762;
    Best Local Similarity 76.9%; Pred. No. 9.2e-91;
    Matches 587; Conservative 0; Mismatches 175; Indels 1; Gaps 1;

QY 46 CCGCGGCTGGCTTGTGTAACCTTCGCTAGTGAAGTTCCAGGGCGAGACCTTACTGCT 105
    |||||
Db 1 CCGCGGCTGGCTTGTGTAACCTTCGCTAGTGAAGTTCCAGGGCGAGACCTTACAAGCT 60
    |||||

QY 106 GTCACCACTGCCCTGAAGACCGGTTACCGTCACTTGGCTGTGCTGGTACTACCTGAAC 165
    |||||
Db 61 GTCATTGGCGCCCTCAAGACCGGTTACCGTCACTTGGCTGTGCTGGTACTACCTGAAC 120
    |||||

QY 166 GAGGGTGAAGTTGTGAGGGTATCCGTAACCTTCTGAAAGGAGAACCCCTCGGTGAAGCGT 225
    |||||
Db 121 GAGGGCGAGGTTGGCCAGCGGCTCCGTAATTCCTCGCCGAGGACACACCGCGCTCAAGCGC 180
    |||||

QY 226 GAGGACATCTTCGTCGACCAAGTCTGGACCACTCCACCGTTATGAGGACGCTCCTC 285
    |||||
Db 181 GAGGATCTTCATTTGACAAAGGTTCTGGAAACCACTCCACGAGCTGAGGAGGTCAAG 240
    |||||

QY 286 TGGTCCATTGACGACTCCCTGAAGCGTCTTGGACTGACGTTGATATGTTCTTCGTT 345
    |||||

```

```

Db 241 TGGTCAATTCGAGAACTCTCTCAAGAACTTCGCGCTTGACTACTTGACCTCTTCTCTGTC 300
    |||||
QY 346 CACTGGCCCAATTGCTGCCGAGAAGAAATGCGCAGGGTGAGCCCAAGATTGGCCCTGACGGC 405
    |||||
Db 301 CACTGGCCCTATGCTGCCGAGAAGGGCGACGACTACCAAGCCCAAGATCGGCCCTGATGGC 360
    |||||

QY 406 AAATACGTCATTTCTCAAGACCTGACCGAGAACCCCGAGCCGACATGCGCGGCTATGGAG 465
    |||||
Db 361 AAGTACGTCATCAAGAGAACCTCACTGAGAACCCCGAGCCGACATGCGCGGCTATGGAG 420
    |||||

QY 466 AAGATTATGAGGATCGCAAGGCCAGGTCCTATGTTGTCTTCAACTGGACCATTCGCGAC 525
    |||||
Db 421 GAGATTATGAGTGGCAAGGCCCGCTGCTCAACTGGACCATTCGCGAC 480
    |||||

QY 526 CTTGAGAGATGTTCCAAAGTTGCGCAAGTCTGCTCAAGTCTGCTCAAGCCACCATGAGATTAC 585
    |||||
Db 481 CTGAAGCAGCTCANGGCCCTTCGCAAGGTCAAGCCCGCTTAACCAAGATCGAGATTAT 540
    |||||

QY 586 CCCTTCTCTGCCCAAGCAGGAGTGTGTGAGTACTGCTTCTTCCAGAAACATTATGCCCGTG 645
    |||||
Db 541 CTTTCTCTGCCCAAGCAGGATCTGCTCAAGTCTGCTCGAGAGGACATCTCTCCCGCT 600
    |||||

QY 646 GCCTACTCTCTCTGGGTGCGAGAACAGAGTTCCACACCGGTGAGCGGGTGAAGGAG 705
    |||||
Db 601 GCTTACTCTCTCCCTCGGTTCCAGAACAGAGTTCCACACTGGTGAAGCCGCTCGGCTCC 660
    |||||

QY 706 AACAGAGCTCTGAACGAGATGCGCGAGAAGGGCGGCAACACCTTGTCTCAGGTTCTTATT 765
    |||||
Db 661 AACAGAGCTCTGAACGAGTGTGCGGACCGGTAGCGGCAACACCTTGTCTCAGGTTCTTATT 720
    |||||

QY 766 GCTTGGGCTGTGCGCGCTGGTGTGCTGCTGCTTCTTCCCAAGAGCT 808
    |||||
Db 721 GCTTGGGCTGTGCGCGCGG-TAGCTGTTCTGCCCAAGAGCT 762
    |||||

RESULT 9
AX472807
LOCUS Sequence 22 from Patent EP1213354.
DEFINITION AX472807
ACCESSION AX472807
VERSION AX472807.1 GI:22207656
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Escherichia coli
REFERENCE 1
AUTHORS Asako,H., Matsumura,K., Shimizu,M., Ito,N. and Wakita,R.
TITLE Process for producing optically active 4-halo-3-hydroxybutanoate
JOURNAL Patent: EP 1213354-A 22 12-JUN-2002;
Sumitomo Chemical Company, Limited (JP)
FEATURES
    source
        1..417
            /organism="Escherichia coli"
            /mol_type="unassigned DNA"
            /db_xref="taxon:562"

ORIGIN
    Query Match 42.6%; Score 417; DB 6; Length 417;
    Best Local Similarity 100.0%; Pred. No. 3.8e-79;
    Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 CCAAGGTGTGGNACCACTCCACCGTTATGAGGAGTCTCTGTGCTTATGAGGAGTCTCC 304
    |||||
Db 1 CCAAGGTGTGGNACCACTCCACCGTTATGAGGAGTCTCTGTGCTTATGAGGAGTCTCC 60
    |||||

QY 305 TGAAGCGTCTTGGACTTGAAGTGTGATATGTTCTCTGTTCACTGGCCCAATTGTCGCG 364
    |||||
Db 61 TGAAGCGTCTTGGACTTGAAGTGTGATATGTTCTCTGTTCACTGGCCCAATTGTCGCG 120
    |||||

QY 365 AGAAGATGGCCAGGTTGAGCCCAAGATTGGCCCTGAGGCAATACGTCATTCTCAAGG 424
    |||||

```


ORIGIN	/number=2	
Query Match	62.1%;	Score 607.6; DB 8; Length 2397;
Best Local Similarity	76.2%;	Pred. No. 3.8e-120;
Matches 789; Conservative	0; Mismatches 189; Indels 58; Gaps 1;	
QY	1	ATGCTTAACGGAAAGACTTTTCAATTTGAGCAACGGCGTCAAGATTCTCTGGCGTGGCTTT 60
DB	693	ATGCTTAGCGGAAAGACCTTCAAGCTCAGCAACGGCGTCAAGATTCTCTGGCGTGGCTTT 752
QY	61	GGTACCTTCGCTAGTGAAGTTTCCAAGGGCGAGACCTATATGCTGTGTCACCACTGCGCTG 120
DB	753	GGTACCTTCGCGAGAGGGCGCTCCCGTGAGACCTACCGCGCGTCAAGAGCGCCCTC 812
QY	121	AAGACCGGTTACCGTCACTTGAAGTGTGCTGTGTTACTACTTGAACGAGGGTGAGTTGGT 180
DB	813	GAGGTCCGATACAGGCACTTGAAGTGTGCTGTGTTTCTACAGACGAGGACGAGTTGGC 872
QY	181	GAGGGTATCCGTGACTTTCCTGAAGAGAAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
DB	873	GATCCGTTTCGGGACTTCTCTCAAGGAGAAACCCCTCTGTCAAGCGCGAGGACATCTTCATC 932
QY	241	TGCACCAAGGTGTGAACCACTCCACCGTTATGAGGACGTCCTCTGTCATTTGACGAC 300
DB	933	TGCACCAAGGTGTGAACCACTTACCGTCCCGAGGACGTTGCTGTGATCGAGGAC 992
QY	301	TCCCTGAAGCGCTTTGGACTTGTACTGTGATATGTTCTCGTTCATGTCGCCCATTTGCT 360
DB	993	TCCCTGAAGAGCTCAAGACCGACTATGTGACCTCTTTCCTTATCCACTGGCCCATTTGCC 1052
QY	361	CGCGAGAAGATGGCCAGGTTGAGCCCAAGATTGGCCCTGACGGCA----- 406
DB	1053	CGCGAGAAGAAAGCCAGGACAAGCTAAGATCGGCCCTGATGAAAGGTTTGTCAATTC 1112
QY	407	-----AATAGCTCATCTCAA 422
DB	1113	TTGCTGCTTGAAGCATAAATGAATCAACTAACCGACTGCTTAGTACGTTATTTTGA 1172
QY	423	GGACTTGACCGAGAACCCCGAGCCACATGCGCGCTATGGAGAAATTTATGAGGATCG 482
DB	1173	GGACTTGACCGAGGACCCCAAGCTACTTTGGCAGCGATGGAGAAATTTATGAGGACAA 1232
QY	483	CAAGGCCAGGTCCATTGGTGTCTCAACTGAGACCATTTGCCGACCTTGAGAAAGATGTCCAA 542
DB	1233	GCTGGCCGCTCCATTGGTGTCTCAACTGAGACCATTTGAAGGCTTCGAGAGCTTCCTCAA 1292
QY	543	GTTTCGCAAGGTTCATGCTTCACGCAACCGAGATTCAGCCCTTCTTGCCCAACGA 602
DB	1293	GTACGCCAAGGTCAAGCTCATGTCAACAGATCGAATCCACCCCTTCTTCCCAACGA 1352
QY	603	GGAGTGTGTGAGTACTGTTCTTCAAGAACATTATGCCGTGGCGCTACTCTCTCTGG 662
DB	1353	GGAGTCTATCCAGTACTGCTGGAAGAACGACATCTGCGCCGAGGCTACTCTCCCTCTGG 1412
QY	663	CTCGCAGAACCAAGTTCCACACCGGTGAGCGGGTCAGCGAGAACAGACTCTGAACGA 722
DB	1413	CTCGCAGAACCAAGTTCCTTACACCGGTGAGCGGTGTACGCGAGAACAGACCTTCAACGA 1472
QY	723	GATCGCGGAGAGGGCGGCAACACCTTGCTCAGGTTCTTATGCTGGGCTCTGCGCCG 782
DB	1473	GATTGCCAGAGGGGGCAACACCTTGGTCAAGTCTCATTTGATGGGGTCTGCGCCG 1532
QY	783	TGGCTACGTGTTTCTCCCAAGAGCTCCACCCCAAGCGCATTTGAGTCCAACTTCAAGAG 842
DB	1533	TGGCTACGTGTTTCTCCCAAGAGCTCCACCCCGCGCATTTGAGTCCAACTTCAAGAG 1592
QY	843	CATTGAGCTCTCCGATGCGGACTTTGAAGCCATCAATGCGCTTGCAGAGGTCGTCACTT 902
DB	1593	CATCGAGCTGTCCGACGAGACTACGAGGCGGTCAACGCTGTGCGCAAGGGCGGTCACTT 1652
QY	903	CCGTTTCTCAACATGAAGGATACTTTTCGGATATGATGTCTGGCCCGAGGAGACCGCCAA 962

Db	1653	CCGCTTCGTCAACATGATGAGGACACACCTTTGGCTACGAGTATGGCCGAGGACCGCCAA	1712			
QY	963	GAACCTGTCTCGTGA	978			
DB	1713	GAACATCTCCGCTAA	1728			
RESULT 7						
LOCUS	AX472800	697 bp	DNA linear PAT 09-AUG-2002			
DEFINITION	Sequence 15 from Patent EP1213354.					
ACCESSION	AX472800					
VERSION	AX472800.1	GI:22207649				
KEYWORDS						
SOURCE	Escherichia coli					
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.					
REFERENCE	1					
AUTHORS	Asako, H., Matsumura, K., Shimizu, M., Ito, N. and Wakita, R.					
TITLE	Process for producing optically active 4-halo-3-hydroxybutanoate					
JOURNAL	Patent: EP 1213354-A 15 12-JUN-2002;					
	Sumitomo Chemical Company, Limited (JP)					
FEATURES	Location/Qualifiers					
source	1..697					
	/organism="Escherichia coli"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:562"					
ORIGIN						
Query Match 60.2%; Score 588.6; DB 6; Length 697;						
Best Local Similarity 94.4%; Pred. No. 5.1e-116;						
Matches 611; Conservative 0; Mismatches 35; Indels 1; Gaps 1;						
QY	4	TCTAAGCGAAGACATTTTCACTTGAAGCAACGGCGTCAAGATTCTCTGGCGTGGCTTTGGT	63			
DB	52	TCCAACGGAANAATTTTCACTGTAGCAACGCGGTCAAAATTTCTGGCGTGGCTTTGGT	111			
QY	64	ACCTTGGCTAGTGAAGTTTCAAGGGCGAGACCTTACTCTGTCTACACTGCCCTCCCTGAAG	123			
DB	112	ACCTTGGCTAGTGAAGTTTCAAGGGCGAAACCTATNCTGTCTACACTGCCCTCCCTGAA	171			
QY	124	ACCGTTACCGTCACTTGGACTGTGCTGTACTACTGAACGAGGGTGAGTTGGTGTAG	183			
DB	172	ACCGTTACCGTCACTTGGACTGTGCTGTACTACTGAACGAGGGTGAGTTGGTGTAG	231			
QY	184	GGTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGCTGC	243			
DB	232	GGTATCCGTGACTTCTCTGAAGGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGCTGC	291			
QY	244	ACCAAGGTGTGAAACCACTCCACCGTTATGAGGACGTCCTCTGGTCCAATTGACGACTCC	303			
DB	292	ACCAAGGTGTGAAACCACTCCACCGTTATGAGGACGTCCTCTGGTCCAATTGACGACTCC	351			
QY	304	CTGAAGGCTTTGAGACTTGAATGATATGTTCTCTGTTACTTGCGCCCAATTCGTCGC	363			
DB	352	CTGAAGGCTTTGAGACTTGAATGATATGTTCTCTGTTACTTGCGCCCAATTCGTCGC	411			
QY	364	GAGAGATGCGCGAGGTTGAGCCCAAGATTGGCCCTGACGCGAAATACGTCATCTCAAG	423			
DB	412	GAGAGATGCGCGAGGTTGAGCCCAAAATTTGGCCCTGACGCGAAATACNTCTTCAAG	471			
QY	424	GACCTGACCGGAGAACCCGAGCCACATGGCGGCTATGGAGAAAGATTTATGAGGATCGC	483			
DB	472	GACCTGACCGGA-AAANCCCNANCCACCTGCGCGCTATGGAAAAAATTTNTGANGATCC	530			
QY	484	AAGGCGAGTCCATTGGTGTCTCCAACTGGACCATTTGCCGACCTTGAGAGATGTCCTCAAG	543			
DB	531	AAGGCGAGTCCATTGGTGTCTCCAACTGGACCATTTGCCGACCTTGAGAGATGTCCTCAAG	590			
QY	544	TTGCGCAAGGTTCATCGCTCACGCCAACAGATTCAGATTACCCCTTCTGCGCCCAACGAG	603			
DB	591	TTGCGCAAGGTTCATCGCTCACGCCAACAGATTCAGATTACCCCTTCTGCGCCCAACGAG	650			

Higashihiroshima, Hiroshima 739-0046, Japan
(E-mail: skao_t@rib.go.jp, Tel: 81-824-20-0824, Fax: 81-824-20-0808)

FEATURES

Location/Qualifiers
1..1183
/organism="Aspergillus oryzae"
/mol_type="mRNA"
/strain="R1B40"
/db_xref="taxon:5062"
/clone="JK1125"
33..1030
/codon_start=1
/product="reductase-like protein"
/protein_id="BAC56099.1"
/db_xref="GI:28144501"
/translation="MSGKFTLSNGVKIPGVGFTFASGSKGTYQAVTKALERYG
RHLDCAFYLNEDVGGIDHDKNPSVEREDIFVCTKVNHHLRPEDEVMSVNSL
KRLDLYVDLHVHPITASEKEQKPIGPDGKYVILKELTENPPTWRAMEKIYED
GKAATGVSNTWITPGLSKLPFAIEIKPHVNOIEIHPFLPNNEVQPCFKNIDILPEAYS
PLGSONQVPTTGERVSENKTLNEIAQKGNLTAQVLIAGLRGRYVVLPKSSNPRIE
SNFKSIELTDDEDFEAVNKAEGRHFVFNKMDTFGYDVPPEETAKNLSA"

ORIGIN

Query Match 75.1%; Score 734.8; DB 8; Length 1183;
Best Local Similarity 84.5%; Pred. No. 1.9e-147;
Matches 826; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 1 ATGTCTAACGGAAGACTTTTCCATGATGAGCAACGCGCTCAAGATTCTCGTCGCGGTTT 60
DB 53 ATGTGCAACGGAAGACTTTTCCATGATGAGCAACGCGCTCAAGATTCTCGTCGCGGTTT 112
QY 61 GGTACTTTCGTAGTGAAGTTTCAAGGGGAGACCTATGCTGTGTCACCTGCGGTTT 120
DB 113 GGTACTTTCGTAGTGAAGTTTCAAGGGGAGACCTATGCTGTGTCACCTGCGGTTT 172
QY 121 AAGACCGGTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 173 GAGACCGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 232
QY 181 GAGGATATCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 233 GATGATATCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 292
QY 241 TGCACCAAGGTGTGAACCACTCCACGCTTATGAGGAGCTCTCTGCTGCTGCTGCTGCT 300
DB 293 TGCACCAAGGTGTGAACCACTCCACGCTTATGAGGAGCTCTCTGCTGCTGCTGCTGCT 352
QY 301 TCCCTGAAGCGTCTTGGACTTGACTGATGATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 353 TCCCTGAAGAGACTCCGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACT 412
QY 361 GCCGAGAGATGCGCAGGTGAGCCCAAGATTGCGCTGAGCCCAATACGTCATCTTC 420
DB 413 TCCGAGAGAGGAGGACGAGGAAACCCCAAGATTGCGCTGAGCCCAATACGTCATCTTC 472
QY 421 AAGACCTTGACGAGAACCCCGAGCCACATGCGGCGCTATGAGAGAGATTATGAGGAT 480
DB 473 AAGAGCTCACTGAGAACCCCGAGCCACATGCGGCGCTATGAGAGAGATTATGAGGAT 532
QY 481 CGAAGGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 533 GCGAAGGCGAGGCGATCGGTGTTTCCAACTGGACCACTCCCGGCTTCGAGAGAGCTGTC 592
QY 541 AAGTTTCGCAAGGTCATCCCTCAGCCCAACAGATTCGAGATTACCCCTTCTGCTGCTGCT 600
DB 593 AAGTTTCGCGAGATCAAGCTCATGTCACACAGATTGAGATCCACCCCTTCTGCTGCTGCT 652
QY 601 GAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 653 AACGAGTTGGTCCAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712
QY 661 GGCTCGCAGACCGAGTTTCCACACCGGTCAGCGGTCAGCGGAGCAAGACTCTGCAAC 720
DB 713 GGCTCCGAGACCGAGTTTCCACACCGGTCAGCGGTCAGCGGAGCAAGACTCTGCAAC 772

QY 721 GAGATCGCGGAGAGCGGCAACACCTTGTCTCAGGTTCTTATTGCTGGGCTGCGC 780
DB 773 GAGATCGCGGAGAGCGGCGCAACACCTTGTCTCAGGTTCTTATTGCTGGGCTGCGC 832
QY 781 CTGCTAGCTGCTGCTTCTCCCAAGAGCTCCCAACCCCAAGCGGATGAGTCCAACTCAAG 840
DB 833 CGCGCTACGTGCTGCTTCTCCCAAGAGCTCCCAACCCCGGCTATCGAGTCCAACTCAAG 892
QY 841 AGCATGAGCTCTCGATGCGGACCTTTGAAGCCATCAATGCGGTTGCCCAAGGCTGCTCAC 900
DB 893 AGCATGAGCTCTCGATGCGGACCTTTGAAGCCCGTCAACAGGTCGCGGCTGCTCAC 952
QY 901 TTCCGTTTCTGCTCAACATGAAGGATGCTTTCGATGATGATGCTGCGCCCGAGAGACCGCC 960
DB 953 TTCCGTTTCTGCTCAACATGAAGGATGCTTTCGATGATGATGCTGCGCCCGAGAGACCGCC 1012
QY 961 AAGAACCTGCTGCTGCTGA 978
DB 1013 AAGAACCTGCTGCTGCTGA 1030

RESULT 6
ANI496025 2397 bp DNA linear PLN 17-JUL-2003
LOCUS Aspergillus nidulans gldB gene for NADP(+) -dependent glycerol
DEFINITION dehydrogenase, exons 1-2.
ACCESSION AJ496025
VERSION AJ496025.1 GI:32879319
KEYWORDS gldB gene; NADP(+) -dependent glycerol dehydrogenase.
SOURCE Emericella nidulans (anamorph: Aspergillus nidulans)
ORGANISM Emericella nidulans
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
REFERENCE 1
AUTHORS de Vries, R.P.
TITLE Glycerol dehydrogenase, encoded by gldB is essential for
osmotolerance in Aspergillus nidulans
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2397)
AUTHORS de Vries, R.P.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2002) de Vries R.P., Microbiology, Utrecht
University, Padualaan 8, Utrecht, 3584 CH, NETHERLANDS
FEATURES
Source 1..2397
/organism="Emericella nidulans"
/mol_type="genomic DNA"
/db_xref="taxon:162425"
693..1728
/gene="gldB"
join(693..1100,1159..1728)
/gene="gldB"
/EC_number="1.1.1.72"
/function="glycerol biosynthesis"
/codon_start=1
/product="NADP(+) -dependent glycerol dehydrogenase"
/protein_id="CAD42649.1"
/db_xref="GI:32879320"
/translation="NSSGKTKLSNGVTIPAVGFTFASGSGTETVAVKALEVGY
RLDCAFYLNEDVGGIDHDKNPSVEREDIFVCTKVNHHLRPEDEVMSVNSL
KRLDLYVDLHVHPITASEKEQKPIGPDGKYVILKELTENPPTWRAMEKIYED
GKAATGVSNTWITPGLSKLPFAIEIKPHVNOIEIHPFLPNNEVQPCFKNIDILPEAYS
PLGSONQVPTTGERVSENKTLNEIAQKGNLTAQVLIAGLRGRYVVLPKSSNPRIE
SNFKSIELTDDEDFEAVNKAEGRHFVFNKMDTFGYDVPPEETAKNLSA"
-693..1100
/gene="gldB"
/number=1
1101..1158
/gene="gldB"
/number=1
1159..1728
/gene="gldB"

gene
CDS
exon
intron
exon

Db 181 GAGGGTATCCGTGACATCTCTGTAAGGAAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
Qy 241 TGCAACAGGTGTGGAAACACCTCCACCGTTATGAGGACGCTCTCTGGTCCATTGACGAC 300
Db 241 TGCACCAAGGTGTGGAAACACCTCCACCGTTATGAGGACGCTCTCTGGTCCATTGACGAC 300
Qy 301 TCCTGAAGCGTCTTGACATCTGACTAGCTGTGATATGTTCTCTGTTCACTGGCCCATTCGCT 360
Db 301 TCCTGAAGCGTCTTGACATCTGACTAGCTGTGATATGTTCTCTGTTCACTGGCCCATTCGCT 360
Qy 361 GCCGAGAAGATGCCAGGCTGAGCCCAAGATTGGCCCTGACGCGCAAAATACGTCATTCTC 420
Db 361 GCCGAGAAGATGCCAGGCTGAGCCCAAGATTGGCCCTGACGCGCAAAATACGTCATTCTC 420
Qy 421 AAGGACCTGACCGAAGAACCCCGAGCCACATGCGCGCTATGAGAAGATTTATGAGGAT 480
Db 421 AAGGACCTGACCGAAGAACCCCGAGCCACATGCGCGCTATGAGAAGATTTATGAGGAT 480
Qy 481 CGCAAGCCAGGTCCATTGCTGCTCAACTGACCAATTCGCCACCTTGAGAGATGTC 540
Db 481 CGCAAGCCAGGTCCATTGCTGCTCAACTGACCAATTCGCCACCTTGAGAGATGTC 540
Qy 541 AAGTTCCCAAGGTCATGCTCAGCCCAACAGATCGAGATTCAACCCCTTCCTGCCCCAAC 600
Db 541 AAGTTCCCAAGGTCATGCTCAGCCCAACAGATCGAGATTCAACCCCTTCCTGCCCCAAC 600
Qy 601 GAGGAGCTGGTGCAGTACTGCTTCCAGAAACATATGCGCGTGGCTACTCTCTCTG 660
Db 601 GAGGAGCTGGTGCAGTACTGCTTCCAGAAACATATGCGCGTGGCTACTCTCTCTG 660
Qy 661 GGCTCGAGAACAGGTTCCACACCGGTGAGCGGTGAGCGGTCAGCGAACAGATCTCGAAC 720
Db 661 GGCTCGAGAACAGGTTCCACACCGGTGAGCGGTGAGCGGTCAGCGAACAGATCTCGAAC 720
Qy 721 GAGATCCCGAAGAGCGGCAACACCTTGTCTCAGGTTCTTATGCTGGGTCTGCGC 780
Db 721 GAGATCCCGAAGAGCGGCAACACCTTGTCTCAGGTTCTTATGCTGGGTCTGCGC 780
Qy 781 CGTGGCTACGTCTCCCAAGAGCTCCAAACCGGATGAGTCCCACTTCAAG 840
Db 781 CGTGGCTACGTCTCCCAAGAGCTCCAAACCGGATGAGTCCCACTTCAAG 840
Qy 841 AGCATTTAGCTCTCCGATGCGGACTTTGAAGCCATCAATCCGTTGCAAGGCTGCTAC 900
Db 841 AGCATTTAGCTCTCCGATGCGGACTTTGAAGCCATCAATCCGTTGCAAGGCTGCTAC 900
Qy 901 TTCGTTTCGTCACATGAAGATCTTTCCGATATGATCTGCGGAGAGCGCC 960
Db 901 TTCGTTTCGTCACATGAAGATCTTTCCGATATGATCTGCGGAGAGAGCGCC 960
Qy 961 AAGAACCTGCTGCGTGA 978
Db 961 AAGAACCTGCTGCGTGA 978

RESULT 3
AX815044
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AX815044
Sequence 4 from Patent EP1323827.
AX815044
AX815044.1 GI:39104167
Penicillium citrinum
Penicillium citrinum
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
Asako,H., Wakita,R. and Itoh,N.
Method for producing optically active
2-hydroxycycloalkane-carboxylic acid ester
Patent: EP 1323827-A 4 02-JUL-2003;
Sumitomo Chemical Company, Limited (JP)
Location/Qualifiers

source 1..978
/organism="Penicillium citrinum"
/mol_type="unassigned DNA"
/db_xref="taxon:5077"
CDS 1..978
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB83559.1"
/db_xref="GI:39104168"
/translation="MSNGKFTLNSGVKIPGVGFTFASGSKGKGTVTATVTTALKTGY
RHLCAWYLNVEGVGEGIDFLKNSPKREDIFVCTKWNHUIHVEDVLSIDSL
KRLGDYVDFMELVHPITAAEKNGGEPKIGDKYVILKDLTENPEPTWAKMEKIYD
RKASIGVSNVTIADLEKMSKFAKMPHANOIEIHPFLNPELVOYCFKSNIMEVAYS
PLGSONQVPTTGERSVSKNTLINEIAEKGNLTAVLJAWLGRGVVLPKSNPKRIE
SNFKSIELSDADFEAINNAVAKGRHFRFVNMDTFGYDVPETAKNLSA"

ORIGIN
Query Match 100.0%; Score 978; DB 6; Length 978;
Best Local Similarity 100.0%; Pred. No. 1.1e-199;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGCGTGGCTTT 60
Db 1 ATGTCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCTCTGGCGTGGCTTT 60
Qy 61 GGTACCTTCGCTAGTGAAGTTTCAAGGGCGAGACCTATCTGCTGCCACCTGCCCTG 120
Db 61 GGTACCTTCGCTAGTGAAGTTTCAAGGGCGAGACCTATCTGCTGCCACCTGCCCTG 120
Qy 121 AAGACCGGTTACCGTCACTTGGACTGTGCTTGGTACTTACCTGAACGAGGGTGAGTTGGT 180
Db 121 AAGACCGGTTACCGTCACTTGGACTGTGCTTGGTACTTACCTGAACGAGGGTGAGTTGGT 180
Qy 181 GAGGATTCGGTACTTCTGAAAGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
Db 181 GAGGATTCGGTACTTCTGAAAGAGAACCCCTCGGTGAAGCGTGAGGACATCTTCGTC 240
Qy 241 TGCACCAAGGTGTGGAACACCTCCACCGTTATGAGGACGCTCTCTGGTCCATTGACGAC 300
Db 241 TGCACCAAGGTGTGGAACACCTCCACCGTTATGAGGACGCTCTCTGGTCCATTGACGAC 300
Qy 301 TCCTGAAGCGTCTTGGACTTACCTGATATGTTCTCTGTTCACTGGCCCATTCGCT 360
Db 301 TCCTGAAGCGTCTTGGACTTACCTGATATGTTCTCTGTTCACTGGCCCATTCGCT 360
Qy 361 GCCGAGAAGATGCGCAGGCTGAGCCCAAGATTGGCCCTGACGCGCAAAATACGTCATTC 420
Db 361 GCCGAGAAGATGCGCAGGCTGAGCCCAAGATTGGCCCTGACGCGCAAAATACGTCATTC 420
Qy 421 AAGGACCTGACCGAAGAACCCCGAGCCACATGCGCGCTATGAGAAGATTTATGAGGAT 480
Db 421 AAGGACCTGACCGAAGAACCCCGAGCCACATGCGCGCTATGAGAAGATTTATGAGGAT 480
Qy 481 CGCAAGCCAGGTCCATTGCTGCTCAACTGACCAATTCGCCACCTTGAGAGATGTC 540
Db 481 CGCAAGCCAGGTCCATTGCTGCTCAACTGACCAATTCGCCACCTTGAGAGATGTC 540
Qy 541 AAGTTCCCAAGGTCATGCTCAGCCCAACAGATCGAGATTCAACCCCTTCCTGCCCCAAC 600
Db 541 AAGTTCCCAAGGTCATGCTCAGCCCAACAGATCGAGATTCAACCCCTTCCTGCCCCAAC 600
Qy 601 GAGGAGCTGGTGCAGTACTGCTTCCAGAAACATATGCGCGTGGCTACTCTCTCTG 660
Db 601 GAGGAGCTGGTGCAGTACTGCTTCCAGAAACATATGCGCGTGGCTACTCTCTCTG 660
Qy 661 GGCTCGCAAGAACAGGTTCCCAACCGGTGAGCGGTGAGCGGTCAGCGAACAGATCTGAAC 720
Db 661 GGCTCGCAAGAACAGGTTCCCAACCGGTGAGCGGTGAGCGGTCAGCGAACAGATCTGAAC 720
Qy 721 GAGATCCCGAAGAGCGGCAACACCTTGTCTCAGGTTCTTATGCTGGGTCTGCGC 780
Db 721 GAGATCCCGAAGAGCGGCAACACCTTGTCTCAGGTTCTTATGCTGGGTCTGCGC 780

Sumitomo Chemical Company, Limited (JP)		721		GAGATCGCGGAGAGGGCGGCAACACCTTGCTCAGGTTCTTATTGCTGGGTCGCGC		780	
FEATURES	Location/Qualifiers	1..978					
	Source						
CDS	/organism="Penicillium citrinum"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:5077"						
	1..978						
	/note="unnamed protein product"						
	/codon_start=1						
	/protein_id="CAD43580.1"						
	/db_xref="GI:22207641"						
	/db_xref="REMTREMBL:CAD43580"						
	/translation="MSGKFTLSNGVKIPGVGFTFASSEKSGEYTVAVTTALKTGY						
	RHLDCAWYLNEGEVGEIRDFLENPSVKREDIFVCTKVNHLHRYEDVLWSDLSL						
	RELGLDYDMFLVHWPIAAEKNGQGEPIKIPDGKYVILKDLTENPPTWRAMEKIYED						
	RKARISGVSNWTIADLEKMSKPAKVMPHANQIEIHPFLPNEELVQVCFSKNIMPVAYS						
	PLGSONOVPTTGERVSENKTLNEIAEKGNLTIAQVLIAGLRRGVVLPKSSNPKEIE						
	SNFKSIELSDADFEAINAVAKGRHFRFVNMKDTFGYDWPETAKNLSA"						
ORIGIN							
Query Match		100.0%; Score 978; DB 6; Length 978;					
Best Local Similarity		100.0%; Pred. No. 1.1e-199;					
Matches		978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60	QY	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60
Db	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60	Db	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60
QY	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120	QY	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120
Db	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120	Db	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120
QY	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180	QY	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180
Db	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180	Db	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180
QY	181	GAGGCTATCCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCGTGAACACATCTTCGTC	240	QY	181	GAGGCTATCCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCGTGAACACATCTTCGTC	240
RESULT 2							
AX472815							
LOCUS		AX472815		978 bp		linear	
DEFINITION		Sequence 30 from Patent EPI213354.					
ACCESSION		AX472815					
VERSION		AX472815.1		GI:22207666			
KEYWORDS		Penicillium citrinum					
SOURCE		Penicillium citrinum					
ORGANISM		Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;					
REFERENCE		1		Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.			
AUTHORS		Asako, H., Matsumura, K., Shimizu, M., Ito, N. and Wakita, R.					
TITLE		Process for producing optically active 4-halo-3-hydroxybutanoate					
JOURNAL		Patent: EP 1213354-A 30 12-JUN-2002;					
		Sumitomo Chemical Company, Limited (JP)					
FEATURES							
Source		1..978					
		/organism="Penicillium citrinum"					
		/mol_type="unassigned DNA"					
		/db_xref="taxon:5077"					
		1..978					
		/note="unnamed protein product"					
		/codon_start=1					
		/protein_id="CAD43583.1"					
		/db_xref="GI:22207667"					
		/db_xref="REMTREMBL:CAD43583"					
		/translation="MSGKFTLSNGVKIPGVGFTFASSEKSGEYTVAVTTALKTGY					
		RHLDCAWYLNEGEVGEIRDFLENPSVKREDIFVCTKVNHLHRYEDVLWSDLSL					
		RELGLDYDMFLVHWPIAAEKNGQGEPIKIPDGKYVILKDLTENPPTWRAMEKIYED					
		RKARISGVSNWTIADLEKMSKPAKVMPHANQIEIHPFLPNEELVQVCFSKNIMPVAYS					
		PLGSONOVPTTGERVSENKTLNEIAEKGNLTIAQVLIAGLRRGVVLPKSSNPKEIE					
		SNFKSIELSDADFEAINAVAKGRHFRFVNMKDTFGYDWPETAKNLSA"					
CDS							
Query Match		100.0%; Score 978; DB 6; Length 978;					
Best Local Similarity		100.0%; Pred. No. 1.1e-199;					
Matches		978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60	QY	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60
Db	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60	Db	1	ATGCTAACGGAAGACTTTCACATTGAGCAACGGCGTCAAGATTCCTGGCGTCGGCTTT	60
QY	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120	QY	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120
Db	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120	Db	61	GGTACCTTCGCTAGTAGAGGTTTCCAAAGGGCGAGACCTATACCTGCTCACCACCTGCCCTG	120
QY	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180	QY	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180
Db	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180	Db	121	AAGACCGGTTACCGTCACTTGGACTGTGCCTGGTACTACCTGAACGAGGTTGAGTTGGT	180
QY	181	GAGGCTATCCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCGTGAACACATCTTCGTC	240	QY	181	GAGGCTATCCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCGTGAACACATCTTCGTC	240
Db	181	GAGGCTATCCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCGTGAACACATCTTCGTC	240	Db	181	GAGGCTATCCGTGACTTCTCTGAAGAGAACCCCTCGTGAACGCGTGAACACATCTTCGTC	240
QY	241	TGCACCAAGGTGTGGAACCACTCCACGGTTATGAGGACGTCCTCTGGTCCATTGAGCAC	300	QY	241	TGCACCAAGGTGTGGAACCACTCCACGGTTATGAGGACGTCCTCTGGTCCATTGAGCAC	300
Db	241	TGCACCAAGGTGTGGAACCACTCCACGGTTATGAGGACGTCCTCTGGTCCATTGAGCAC	300	Db	241	TGCACCAAGGTGTGGAACCACTCCACGGTTATGAGGACGTCCTCTGGTCCATTGAGCAC	300
QY	301	TCCCTGAAGCGCTTGGACTTGACTACGTTGATATGTTCTCTGTTCACTGCCCATTTGCT	360	QY	301	TCCCTGAAGCGCTTGGACTTGACTACGTTGATATGTTCTCTGTTCACTGCCCATTTGCT	360
Db	301	TCCCTGAAGCGCTTGGACTTGACTACGTTGATATGTTCTCTGTTCACTGCCCATTTGCT	360	Db	301	TCCCTGAAGCGCTTGGACTTGACTACGTTGATATGTTCTCTGTTCACTGCCCATTTGCT	360
QY	361	CCGAGAGAATATGCGCAGGGTGAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC	420	QY	361	CCGAGAGAATATGCGCAGGGTGAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC	420
Db	361	CCGAGAGAATATGCGCAGGGTGAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC	420	Db	361	CCGAGAGAATATGCGCAGGGTGAGCCCAAGATTGGCCCTGACGGCAAAATACGTCATTCTC	420
QY	421	AAGACCTGACGAGAACCCCGACCCACATGGCGCGCTATGGAGAGAGATTATGAGGAT	480	QY	421	AAGACCTGACGAGAACCCCGACCCACATGGCGCGCTATGGAGAGAGATTATGAGGAT	480
Db	421	AAGACCTGACGAGAACCCCGACCCACATGGCGCGCTATGGAGAGAGATTATGAGGAT	480	Db	421	AAGACCTGACGAGAACCCCGACCCACATGGCGCGCTATGGAGAGAGATTATGAGGAT	480
QY	481	CGAAGCCAGGTCATTGTTGTCTCCAACTGACCACTTGGCCGCTTACCGCTTCTCCCTCTG	540	QY	481	CGAAGCCAGGTCATTGTTGTCTCCAACTGACCACTTGGCCGCTTACCGCTTCTCCCTCTG	540
Db	481	CGAAGCCAGGTCATTGTTGTCTCCAACTGACCACTTGGCCGCTTACCGCTTCTCCCTCTG	540	Db	481	CGAAGCCAGGTCATTGTTGTCTCCAACTGACCACTTGGCCGCTTACCGCTTCTCCCTCTG	540
QY	541	AAGTTCCGCAAGGTCATGCTCAGCCCAACAGATCGAGATTACCGCTTCTCCCTCTG	600	QY	541	AAGTTCCGCAAGGTCATGCTCAGCCCAACAGATCGAGATTACCGCTTCTCCCTCTG	600
Db	541	AAGTTCCGCAAGGTCATGCTCAGCCCAACAGATCGAGATTACCGCTTCTCCCTCTG	600	Db	541	AAGTTCCGCAAGGTCATGCTCAGCCCAACAGATCGAGATTACCGCTTCTCCCTCTG	600
QY	601	GAGGAGCTGGTGCAGTACTGTTCTCCAAAGAACATTATGCCCGTGGCTACTCTCCCTCTG	660	QY	601	GAGGAGCTGGTGCAGTACTGTTCTCCAAAGAACATTATGCCCGTGGCTACTCTCCCTCTG	660
Db	601	GAGGAGCTGGTGCAGTACTGTTCTCCAAAGAACATTATGCCCGTGGCTACTCTCCCTCTG	660	Db	601	GAGGAGCTGGTGCAGTACTGTTCTCCAAAGAACATTATGCCCGTGGCTACTCTCCCTCTG	660
QY	661	GGCTCGAGAACACAGGTTCCACACACCGGTGAGCGGGTACGCGAGAACAGACTCTGAAC	720	QY	661	GGCTCGAGAACACAGGTTCCACACACCGGTGAGCGGGTACGCGAGAACAGACTCTGAAC	720
Db	661	GGCTCGAGAACACAGGTTCCACACACCGGTGAGCGGGTACGCGAGAACAGACTCTGAAC	720	Db	661	GGCTCGAGAACACAGGTTCCACACACCGGTGAGCGGGTACGCGAGAACAGACTCTGAAC	720
QY	721	GAGATCGCCGAGAGAGGGGGGCAACACCCCTTCTCAGGTTCTTATTGCTGGGTCGCGC	780	QY	721	GAGATCGCCGAGAGAGGGGGGCAACACCCCTTCTCAGGTTCTTATTGCTGGGTCGCGC	780

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2004, 11:12:09 ; Search time 3874 Seconds
(without alignments)
10942.046 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978

Sequence: 1 atgtctaacgaagaagacttt.....ccaagaacctgtctgcgtga 978

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.mu.*

21: em.ov.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.man.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	978	100.0	978	6	AX472787	AX472787 Sequence
2	978	100.0	978	6	AX472815	AX472815 Sequence
3	978	100.0	978	6	AX815044	AX815044 Sequence
4	978	100.0	996	6	AX472813	AX472813 Sequence
5	734.8	75.1	1183	8	AB079030	AB079030 Aspergill
6	607.6	62.1	2397	8	ANI495025	ANI495025 Aspergill
7	588.6	60.2	597	6	AX472800	AX472800 Sequence
8	471	48.2	762	8	AB014493	AB014493 Gibberell
9	417	42.6	417	6	AX472807	AX472807 Sequence
10	412	42.1	743	6	AX472804	AX472804 Sequence
11	265	27.1	331	6	AX472803	AX472803 Sequence
12	247	25.3	720	8	CNS01C4M	AL112097 Botrytis
13	233.2	23.8	720	8	CNS01SU1	AL112097 Botrytis
14	230.6	23.6	660	8	CNS01BYU	AL114862 Botrytis
15	226.2	23.1	780	8	CNS01BQ2	AL114846 Botrytis
16	208	21.3	720	8	CNS01AFA	AL112862 Botrytis
17	199.6	20.4	1055	6	E11242	E11242 cDNA encodi
18	188.8	19.3	720	8	CNS01DNY	AL117662 Botrytis
19	180.2	18.4	939	8	AY033888	AY033888 Magnapor
20	178.8	18.3	1123	3	BT003280	BT003280 Drosophi
21	174.8	17.9	780	8	CNS01DM8	AL117000 Botrytis
22	144.2	14.7	1121	6	AX470403	AX470403 Sequence
23	144.2	14.7	1121	6	BD177581	BD177581 Erythro
24	144.2	14.7	1121	6	BD177586	BD177586 Erythro
25	142.6	14.6	1077	6	AX470402	AX470402 Sequence
26	142.6	14.6	1077	6	BD177580	BD177580 Erythro
27	142.6	14.6	1077	6	BD177585	BD177585 Erythro
28	141.6	14.5	456	8	CNS01B1R	AL113671 Botrytis
29	137.4	14.0	2359	4	BOVALRDB	M59754 Bovine lens
30	126	12.9	970	4	FIGALDRD	L14950 Sus scrofa
31	126	12.9	970	4	SSU46065	U46065 Sus scrofa
32	124.8	12.8	1156	4	BOVALDRD	M31463 Bovine aldo
33	124.8	12.8	1187	4	S54973	S54973 20 alpha-hy
34	122.2	12.5	456	8	CNS01BBK	AL114024 Botrytis
35	121.2	12.4	2309	8	SSU26463	U26463 Sporidibol
36	121	12.4	1395	10	BC051128	BC051128 Mus muscu
37	119.8	12.2	1310	4	OCU12316	U12316 Oryctolagus
38	119.8	12.2	1494	4	RABALRE	M32818 Rabbit (clo
39	119.2	12.2	1335	6	AR380465	AR380465 Sequence
40	119.2	12.2	1335	9	HUMARXA	M34720 Human aldos
41	119.2	12.2	1354	9	HUMALRMB	J05474 Human aldos
42	119.2	12.2	1355	9	AF328729	AF328729 Homo sapi
43	119.2	12.2	1367	6	AX440479	AX440479 Sequence
44	119.2	12.2	1367	9	HSALRE	X15414 Human mRNA
45	119.2	12.2	1368	6	AX451813	AX451813 Sequence

ALIGNMENTS

RESULT 1	AX472787	978 bp	DNA	linear	PAT 09-AUG-2002
LOCUS	Sequence 2 from Patent EP1213354.				
DEFINITION	AX472787				
ACCESSION	AX472787.1	GI:22207640			
VERSION	AX472787.1				
KEYWORDS	Penicillium citrinum				
SOURCE	Penicillium citrinum				
ORGANISM	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.				
REFERENCE	1				
AUTHORS	Asako,H., Matsumura,K., Shimizu,M., Ito,N. and Wakita,R.				
TITLE	Process for producing optically active 4-halo-3-hydroxybutanoate				
JOURNAL	Patent: EP 1213354-A 2 12-JUN-2002;				

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.478C14
;; CURRENT APPLICATION NUMBER: US/09/702,705
;; CURRENT FILING DATE: 2000-10-30
;; NUMBER OF SEQ ID NOS: 1833
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 340
;; LENGTH: 316
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-702-705-340

Query Match 28.1%; Score 487; DB 4; Length 316;
Best Local Similarity 36.6%; Pred. No. 8.8e-44;
Matches 118; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

QY 6 TFT-LSNGVKIPGVGFGTFASEGSKGETYATVTTALKTGVRHLDCAWYVNLNVEGVEGIR 64
DB 3 TFVELSTKAKMPVIGLGTWKS--PLGKVEAVKVAIDAGYRHIDCAVYVQNEHEVGEAIQ 60

QY 65 DFLKENPSVKREDIFVCTKVMNHLHRYEDVLWSIDDSLKRLGLDYDVMFLVHWPAAEKN 124
DB 61 EKIQEK-AVKREDLFIVSKLWPTFFERPLVKAPEKTLKDLKSLVDVLIHWPQGFSG 119

QY 125 GQGEPK-----IGPDGKYVILKLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEK 178
DB 120 DDLFPKDDKGNAG--GKATFL-----DAWEAMEELVDGLVKALGVSNFHFQIEK 169

QY 179 M--SKFAKVMPHANQIEHPFLPNBELVOYCFSKNIMPVAYSPGSONQVPTTGERVS-- 234
DB 170 LLNKPGLKYKPVTVNQECPHYLTQEKLIQYCHSKGITVTAYSPGSDPRPWAKPDPSSL 229

QY 235 ENKTLNEIAEKGNTLAQVLIAGLRRGYVWLPKSSNPKRIESNFKSIELSDADFEAINA 294
DB 230 EDPKIEIAAKHKHTAAQVLIHFHQRNVIVIPKSVTPARIVENTQVDFKLSDEEMATI 289

QY 295 VAKGRHFRFVNKMDTFGYDWP 316
DB 290 LSFNRNRACNVLSQSHLEDYP 311

RESULT 14
US-09-736-457-340
; Sequence 340, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 340
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-340

Query Match 28.1%; Score 487; DB 4; Length 316;
Best Local Similarity 36.6%; Pred. No. 8.8e-44;
Matches 118; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

QY 6 TFT-LSNGVKIPGVGFGTFASEGSKGETYATVTTALKTGVRHLDCAWYVNLNVEGVEGIR 64
DB 3 TFVELSTKAKMPVIGLGTWKS--PLGKVEAVKVAIDAGYRHIDCAVYVQNEHEVGEAIQ 60

QY 65 DFLKENPSVKREDIFVCTKVMNHLHRYEDVLWSIDDSLKRLGLDYDVMFLVHWPAAEKN 124
DB 61 EKIQEK-AVKREDLFIVSKLWPTFFERPLVKAPEKTLKDLKSLVDVLIHWPQGFSG 119

QY 125 GQGEPK-----IGPDGKYVILKLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEK 178
DB 120 DDLFPKDDKGNAG--GKATFL-----DAWEAMEELVDGLVKALGVSNFHFQIEK 169

QY 179 M--SKFAKVMPHANQIEHPFLPNBELVOYCFSKNIMPVAYSPGSONQVPTTGERVS-- 234
DB 170 LLNKPGLKYKPVTVNQECPHYLTQEKLIQYCHSKGITVTAYSPGSDPRPWAKPDPSSL 229

QY 235 ENKTLNEIAEKGNTLAQVLIAGLRRGYVWLPKSSNPKRIESNFKSIELSDADFEAINA 294
DB 230 EDPKIEIAAKHKHTAAQVLIHFHQRNVIVIPKSVTPARIVENTQVDFKLSDEEMATI 289

QY 295 VAKGRHFRFVNKMDTFGYDWP 316
DB 290 LSFNRNRACNVLSQSHLEDYP 311

RESULT 15
US-09-614-124B-340
; Sequence 340, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 340
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-340

Query Match 28.1%; Score 487; DB 4; Length 316;
Best Local Similarity 36.6%; Pred. No. 8.8e-44;
Matches 118; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

QY 6 TFT-LSNGVKIPGVGFGTFASEGSKGETYATVTTALKTGVRHLDCAWYVNLNVEGVEGIR 64
DB 3 TFVELSTKAKMPVIGLGTWKS--PLGKVEAVKVAIDAGYRHIDCAVYVQNEHEVGEAIQ 60

QY 65 DFLKENPSVKREDIFVCTKVMNHLHRYEDVLWSIDDSLKRLGLDYDVMFLVHWPAAEKN 124
DB 61 EKIQEK-AVKREDLFIVSKLWPTFFERPLVKAPEKTLKDLKSLVDVLIHWPQGFSG 119

QY 125 GQGEPK-----IGPDGKYVILKLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEK 178
DB 120 DDLFPKDDKGNAG--GKATFL-----DAWEAMEELVDGLVKALGVSNFHFQIEK 169

QY 179 M--SKFAKVMPHANQIEHPFLPNBELVOYCFSKNIMPVAYSPGSONQVPTTGERVS-- 234
DB 170 LLNKPGLKYKPVTVNQECPHYLTQEKLIQYCHSKGITVTAYSPGSDPRPWAKPDPSSL 229

QY 235 ENKTLNEIAEKGNTLAQVLIAGLRRGYVWLPKSSNPKRIESNFKSIELSDADFEAINA 294

Best Local Similarity 39.4%; Pred. No. 4e-45;
Matches 121; Conservative 51; Mismatches 106; Indels 29; Gaps 10;

QY 9 LNSGVKIPGVGTFASGSGKETVAVTTALKTGYRHLDCAWYLNNEGEVGEIRD 65
DB 3 TUVLSGGVEMPVIGLGLWRLE--KDELKEVINAIKGYRHFDCAAHYKSEADVGEALAE 60
QY 66 FLKENPSVKREDIFVCTKWNHLYEDVLSIDSLKRLGLDYVDMFLVHWPFAAEKNG 125
DB 61 AFKTG-LVKREBELFTTKIMNSDGH--VBEACKNSLEKLIQIDYLDLYLVHVPMPTRHNA 117
QY 126 QGEPK--TGPCKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMSKFA 183
DB 118 ICKTASLGED--KVLIDIVTISLQOTWEGMEKTVSLGLVRSIGLSYVEFLTRDCLAYS 175
QY 184 KVMPHANOIEIHPFLNPEELVOYCFCKNIMPVAYSPLGSONQVPTTGERVS--ENKTLNE 241
DB 176 KIKPAVSQFETHPYFQDRLSVKFCMKHGVLPHTAHTPLGGAANKDMFGSVSPLDDPVLND 235
QY 242 IAEKGNVLAQVLTAWGLRRGVVLPKSSNPKRIESNFKSIE--LSADADFEAINAVAKGR 299
DB 236 VAKYKGSVAQICLRWGQRTAVIPKSSKTQRLKENLEVLEFQLSDEDMQLIYSI--DR 293
QY 300 HRFVNMKDTFGYDW 315
DB 294 KYRTSLPKXTWGLDY 309

RESULT 9
US-08-801-344-4
Sequence 4, Application US/08801344
Patent No. 6087140
GENERAL INFORMATION:
APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.
TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801.344
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 09820.037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat lens aldose reductase
US-08-801-344-4

Query Match 28.9%; Score 499.5; DB 3; Length 316;

RESULT 10
US-09-498-599-4
Sequence 4, Application US/09498599
Patent No. 6303352
GENERAL INFORMATION:
APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.
TITLE OF INVENTION: MICROBIAL PRODUCTION OF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,599
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 09820.037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat lens aldose reductase
US-09-498-599-4

RESULT 7
US-08-731-320B-4
; Sequence 4, Application US/08731320B
; Patent No. 6416985
; GENERAL INFORMATION:
; APPLICANT: Wayne D. Loeschner,
; APPLICANT: John D. Everard
; APPLICANT: Rebecca Grumet
; TITLE OF INVENTION: DNA Encoding Mannose
; TITLE OF INVENTION: 6-Phosphate Reductase
; TITLE OF INVENTION: and Recombinants
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch,
; MEDIUM TYPE: 360 kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731.320B
; FILING DATE: October 15, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION:
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: apple
; IMMEDIATE SOURCE:
; LIBRARY: N/A
US-08-731-320B-4
Query Match 29.0%; Score 501.5; DB 4; Length 310;
Best Local Similarity 34.8%; Pred. No. 2.3e-45;
Matches 110; Conservative 73; Mismatches 118; Indels 15; Gaps 8;
QY 6 TPTLSNGVKIPGVGFGTFASSGKGTAVTTALKTGYRHLDCAWYVYLNAGEVGEIRD 65
DB 3 TWTLSGVEYMPVIGUGLWRLE--KDELKEVILNAIKIGYRHFDCAAHYKSEADVEALAE 60
QY 66 FLKENPVSVKREDIFVCTKVVNHLRYEDVLSIDSLKRLGLDVDMFLVHWPIAEKNG 125
DB 61 AFKGTG-LVKRELEFITTITKMSDGHG--VWEACKNSLEKLQIDYLDLYLVHYPMPTRHNA 117
QY 126 QGEPK--IGPGKVILKDLTENPEPTWRAMEKIVEDRKARSIGVSNWTIADLEKMSKFA 183
DB 118 IGTASLLGED--KVLDDIVTISLQQTWEGNEKTVSLGLVRSIGLSNYELPLTRDCLAYS 175

QY 184 KVMPHANOIEIHPFLPNEELVQYCFSKNIMPVAYSPQSQOVPTTGBRVS--ENKTLNE 241
DB 176 KIRPAVSQFETHPYFQRDSLVRCMKHGVLPFAHTPLGGAANKDMFGSVPLDDFVLND 235
QY 242 IAEKGCNTLAQVLIAMGLRRGVVLPKSNPKRIESFKSIB--LSDADFEAINAVAKGR 299
DB 236 VAKYKGKVAQICLRWGIQKTAVIPKSKIQRLKENLEVLEFQLSDDEMQLIYSI--DR 293
QY 300 HRFVNMKDTFGYDWW 315
DB 294 KYRTSLPSKTIWGLDVI 309
RESULT 8
US-09-166-438-4
; Sequence 4, Application US/09166438
; Patent No. 6495317
; GENERAL INFORMATION:
; APPLICANT: Wayne D. Loeschner,
; APPLICANT: John D. Everard
; APPLICANT: Rebecca Grumet
; TITLE OF INVENTION: DNA Encoding Mannose
; TITLE OF INVENTION: 6-Phosphate Reductase
; TITLE OF INVENTION: and Recombinants
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch,
; MEDIUM TYPE: 360 kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/166,438
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/731,320
; FILING DATE: October 15, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION:
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: apple
; IMMEDIATE SOURCE:
; LIBRARY: N/A
US-09-166-438-4
Query Match 29.0%; Score 501.5; DB 4; Length 310;
Best Local Similarity 34.8%; Pred. No. 2.3e-45;
Matches 110; Conservative 73; Mismatches 118; Indels 15; Gaps 8;

Db 237 KYGKTPAQLVRLWGLQRTWVTPKTSKVERQENFDVFGFDISGEDMERKAI--DRKYR 294
QY 303 FVNMKDTFGYDWW 315
Db 295 TNQPAKFWGIDLY 307

RESULT 5
US-09-347-803-14
; Sequence 14, Application US/09347803
; Patent No. 6274379
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes

; FILE REFERENCE: BB-1176
; CURRENT APPLICATION NUMBER: US/09/347.803
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092.952
; EARLIER FILING DATE: July 15, 1998
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Glycine max
US-09-347-803-14

Query Match 29.4%; Score 509.5; DB 3; Length 309;
Best Local Similarity 37.6%; Pred. No. 3.2e-46;
Matches 118; Conservative 59; Mismatches 122; Indels 15; Gaps 8;

QY 8 TLSNGVKIPGVGFGTFASGSGKGETYTAVTALKTGYRHLDCAWYLYNEGEVSGIRDFL 67
Db 4 TLNNGFKMPIGLGWRNEG--EIRDLILNSIKIGYHFDCADYKNEAEVGDALKEAF 61
QY 68 KENPSVKREDIFVCTKYVNHLYHRYEDVLSIDSLKRLGIDYDMFLVHWPPIAAEKNG 127
Db 62 -DSGLVKREDIFITTKLWNSDQGH--VLEACKDSLKKLQTLTYLDLYLVHFPVAVRHTGVG 118
QY 128 EPK--IGPDGKYVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMSKPAK 185
Db 119 NTSSPLGDDG--VLDDITTSLETTWHAEDLVSSGLVRSIGISNYDIFLTRDCLAYSKI 176
QY 186 MPHANQIEIHFPFLNEELVQYCFSKNIMPVAYSPGSGNQVPTTGERVS--ENKTLNEIA 243
Db 177 KPANVQIETHPYFQDSDSLVKFCQKHGICVTAHTPLGGAANAENWFGTVSCLDQVLKGLA 236
QY 244 EKGNTLAQVLIANGRLRGYVLPKSSNPKEISNFK--SIELSDADFEAINAVAKGRHF 301
Db 237 EYKKTAAQISLRWGIQRTWVTPKSKLERLKENFQVDFPELSKEDMELIGSI--DRKY 294
QY 302 RFVNMKDTFGYDWW 315
Db 295 RTNQPAVFWGIDLY 308

RESULT 6
US-09-166-412-4
; Sequence 4, Application US/09166412
; Patent No. 6133504
; GENERAL INFORMATION:
; APPLICANT: Wayne D. Loescher,
; APPLICANT: John D. Everard
; APPLICANT: Rebecca Grumet
; TITLE OF INVENTION: DNA Encoding Mannose
; TITLE OF INVENTION: 6-Phosphate Reductase
; TITLE OF INVENTION: and Recombinants
; TITLE OF INVENTION: Produced Therefrom
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch,
; MEDIUM TYPE: 360 kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; OPERATING SYSTEM: (version 3.3)
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/166,412
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/731,320
; FILING DATE: October 15, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION:
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: apple
; IMMEDIATE SOURCE:
; LIBRARY: N/A
; US-09-166-412-4
Query Match 29.0%; Score 501.5; DB 3; Length 310;
Best Local Similarity 34.8%; Pred. No. 2.3e-45;
Matches 110; Conservative 73; Mismatches 118; Indels 15; Gaps 8;
QY 6 TFLNSGVKIPGVGTFASGSGKGETYTAVTALKTGYRHLDCAWYLYNEGEVSGIRD 65
Db 3 TVTLSSGYEMPVIGLGLWRLE--KDELKEVILNAIKIGYRHFDCAAHYKSEADVGEALAE 60
QY 66 FLKENPSVKREDIFVCTKYVNHLYHRYEDVLSIDSLKRLGIDYDMFLVHWPPIAAEKNG 125
Db 61 APTKG-LVKGRELFITTKLWNSDGH--VVEACKSLKQLIDYLDLYLVHYPMTKINA 117
QY 126 QGEPK--IGPDGKYVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMSKPA 183
Db 118 IGTASLLGED--KVLDDITVTSLOQTWEGMEKTVSLGLVRSIGLSNVELFLTRDCLAYS 175
QY 184 KVMPHANQIEIHFPFLNEELVQYCFSKNIMPVAYSPGSGNQVPTTGERVS--ENKTLNE 241
Db 176 KIKPAVSQFETHPYFQDSDSLVKFCMKHGVLPFAHTPLGGAANAANKDMFGSVPLDDFVLND 235
QY 242 IAEKGGNTLAQVLIANGRLRGYVLPKSSNPKEISNFKSIE--LSDADFEAINAVAKGR 299
Db 236 VAKYKGSVAQICLRWGIQRTAVIPKSKIQRLKENLEVLFQSLSDQMQLIYSI--DR 293
QY 300 HRFVNMKDTFGYDWW 315
Db 294 KYRTSLPSKTKGLDLY 309

```

; Sequence 6, Application US/09347803
; Patent No. 6274379
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
; FILE REFERENCE: BB-1176
; CURRENT APPLICATION NUMBER: US/09/347,803
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,952
; EARLIER FILING DATE: July 15, 1998
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Glycine max
; 09-347-803-6

Query Match          31.4%; Score 543.5; DB 3; Length 313;
Best Local Similarity 37.5%; Pred. No. 7.2e-50;
Matches 123; Conservative 62; Mismatches 108; Indels 35; Gaps 9

Qy      5  KFTFLSNGVKIPGUCFCGFASEGSGKGTYYATVTTALKTGYRHLDCAWYVLNGEVGEGR 64
Db      6  KFEFLNTGAKIPVSLGVLQWQAE--PGVAKAVTTAILVGYRHDCQAQYNNQAEIGSALK 63

Qy      65  DFLKENPSVKREDIFVCTKVNHNHLEHVEDVLWSIDDSLKELGLDYVDMFLVHPFAAEKN 124
Db      64  K-LFDDGVVKREDLWITSKLCWSDHASEDVPKALDKTQLDLDLYLHWPVRMKSQ 122

Qy      125  GQGEKPIGPDGVYVILKOLTENPE--PTWRAMEKIYEDRKARSGIYGNWTIADLEKMSKF 182
Db      123  SVG-----FKSYLDQDPDIPFTWKAMEALYDSGKARAIGVSNFSSKKQLQDLMNI 171

Qy      183  AKYMPHANOIEIHPFLPNEELVQVCFKSNIMPVAYSPLGSONQVPTTGERVSENKTLNEI 242
Db      172  ARVPPAVNQVELHPGQQPKLHAFCESGKGVHLSCYSPLGSPGVLKSD---ILKNPWVIEI 228

Qy      243  AEKGGNTLAQVLIANGLRGVYVLPKSNPKRLESNEFKSIELSDAD-----FEAT--NAV 295
Db      229  AEKLGKTPAQVAIRMGLOTHGSHVLPKSTNIESRIKGNDFVDSWISPEVMDKFSKIQDRL 288

Qy      296  AKGRHFRFVNMDKTFG-----YDVWPEE 318
Db      289  IKGTFP----VDETYGAFKTVBELWDGE 312

```

RESULT 3
US-09-347-803-25
US-09-347-803-25
; Sequence 25, Application US/09347803
; Patent No. 6274379
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: O-zoco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
; FILE REFERENCE: BB-1176
; CURRENT APPLICATION NUMBER: US/09/347,803
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,952
; EARLIER FILING DATE: July 15, 1998
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Medicago sativa
US-09-347-803-25

```

Query Match          30.2%; Score 523.5; DB 3; Length 313;
Best Local Similarity 37.2%; Pred. No. le-47;
Matches 123; Conservative 61; Mismatches 106; Indels 41; Gaps 10

Qy      5 KFTFLSNGVKIPGVGFGTFASGSGKGTYTAVTTALTATGYRHLDCAWYVNLGEVGEGR 64
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      6 KFFQLNTGAKIPSVGLGTWQAE--PGVVAKAVTTAVQVGRYHIDCAEAYKNQSEIGSALK 63

Qy     65 DELKENPSVKREDIFVCTKVNNHLYRVEDVLWSIDDSLKELGLDYDVMFVHWPIAAEK 124
       :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db     64 K-LCEDGVVYKEELWITSLKMSDHPEDVPKALDKTLNDLQDLQDLYLIHWPVSM-KR 121

Qy    125 GOGEPKIGPDGKYVILKDILTENPE-----PTWRAMEKIYEDRKARSIGYSNWTIADLEKM 179
       :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db    122 GTIGE-----FMGENLHDAPISPTWKALGALYDSGKAALGSYNFSTKKLQDL 168

Qy    180 SFPAKVMPHANOIEHPLEPNEELVQCFCFSKNIMPVAYSPLGSONQVPTTGERSVENSKIL 239
       :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db    169 LDVARPPAVNQVELHPWCQAOKLHAFCESGKHLSGYSPLGSPGVCLKSD---ILKNPVV 225

Qy    240 NRIAEKGGNTLAQVLIAMGLRGYVWLFPKSNPKRIESNFKEISLSDAD-----FEAI-- 292
       :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db    226 KSLAEKLCKTKPCQVALRWGLAGHSVLFPKSTNEARIKKNLDVVDWSIPEDLPFKFSEIKQ 285

Qy    293 NAVAKGRHRFRVNMMKDTFG-----YDVWPEE 318
       :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db    286 DXLIKRTGF---VNDITYGAFTIEELMDGE 312


RESULT 4
US-09-347-803-10
; Sequence 10, Application US/09347803
; Patent No. 6274379
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
; FILE REFERENCE: BS-1176
; CURRENT APPLICATION NUMBER: US/09/347,803
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,952
; EARLIER FILING DATE: July 15, 1998
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Zea mays
US-09-347-803-10
```

Query Match	30.0%;	Score 518.5;	DB 3;	Length 308;
Best Local Similarity	37.7%;	Pred. No. 3.4e-47;		
Matches 118;	Conservative 60;	Mismatches 120;	Indels 15;	Gaps 8
Qy	9	LSGNVKIPGVGGTGFASGSGKGTYYTAVTALTGTGRVHLDCAMYYILNEGEVGEIGIRFLK	68	
Db	4	LSGHERMAVLGVWRME--KADIRGLIHIALRVGYRHLDCAADYQNEAEVGDALAEAF-	60	
Qy	69	ENPSVKRREDIFVCTTKWNHLHRYEDVILWSDDSLKRLGLDYVDMELVHPPIAAEKNGQG-	127	
Db	61	QTGLVKREDLFITTKLWNSDGH--VLEACKDSKLQLDYLDLYLIHPVPVATRHGTGVGT	118	
Qy	128	-SPKIGPGDKVYILKDLTENPEPTWRAMEKIEYDRKARSIGVSNWTIADLEKNMSKFAKVM	186	
Db	119	TSSALGDDG--VLDDITISLETTWHAMEELVSNGLVRSIGISNYDIFLTRDCLAYAKIK	176	
Qy	187	PHANQIEHPPLNEELVQYCFSKNIPVAVSPILGSONQVPTTGERVS--ENKTLNEIAE	244	
Db	177	PAWNIQETHYPQRDSLKYFCQKHGICVTAHTJGGGTANAEMFGTGVCLDDDPVILKSLAD	236	
Qy	245	KGNGTLAQLVIAWGLRRGVVYLPKSSNPKEISNFK--SIELSDAEPFAINAVAKGRHFR	302	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 02:50:06 ; Search time 45 Seconds
(without alignments)
372.854 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1731

Sequence: 1 MSNGKTTLSNGVKIPGVGF.....MKDTFGYDVWPEBETAKNLSA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	547.5	31.6	344	4	US-09-630-983A-20
2	543.5	31.4	313	3	US-09-347-803-6
3	523.5	30.2	313	3	US-09-347-803-25
4	518.5	30.0	308	3	US-09-347-803-10
5	509.5	29.4	309	3	US-09-347-803-14
6	501.5	29.0	310	3	US-09-166-412-4
7	501.5	29.0	310	4	US-08-731-320B-4
8	501.5	29.0	310	4	US-09-166-412-4
9	499.5	28.9	316	3	US-08-801-344-4
10	499.5	28.9	316	4	US-09-498-599-4
11	498	28.8	298	4	US-09-134-000C-6059
12	494.5	28.6	308	3	US-09-347-803-12
13	487	28.1	316	4	US-09-702-705-340
14	487	28.1	316	4	US-09-736-457-340
15	487	28.1	316	4	US-09-614-124B-340
16	487	28.1	316	4	US-09-589-184-340
17	487	28.1	316	4	US-09-166-412-2
18	486.5	28.1	309	3	US-09-347-803-26
19	486.5	28.1	309	3	US-08-731-320B-2
20	486.5	28.1	309	4	US-09-166-412-2
21	486.5	28.1	309	4	US-09-166-412-2
22	482.5	27.9	308	3	US-09-347-803-8
23	472.5	27.3	316	1	US-08-585-595-3
24	466	26.9	291	4	US-09-107-532A-5063
25	456	26.3	323	4	US-09-702-705-1821
26	456	26.3	323	4	US-09-736-457-1821
27	456	26.3	323	4	US-09-671-325-1821

28	447.5	25.9	323	3	US-08-532-896-2	Sequence 29, Appli
29	445	25.7	327	4	US-09-294-894-29	Sequence 29, Appli
30	438.5	25.3	281	4	US-09-134-001C-5454	Sequence 5454, Ap
31	432	25.0	318	3	US-09-347-803-4	Sequence 4, Appli
32	432	25.0	320	3	US-09-347-803-23	Sequence 23, Appli
33	431	24.9	282	4	US-09-134-001C-5083	Sequence 5083, Ap
34	430.5	24.9	323	3	US-08-853-839-2	Sequence 2, Appli
35	430	24.8	273	4	US-09-684-385-10	Sequence 10, Appli
36	428	24.7	319	3	US-09-347-803-24	Sequence 24, Appli
37	423.5	24.5	364	4	US-09-643-597-172	Sequence 172, App
38	423.5	24.5	364	4	US-09-480-884A-172	Sequence 172, App
39	423.5	24.5	364	4	US-09-702-705-783	Sequence 783, App
40	423.5	24.5	364	4	US-09-736-457-783	Sequence 783, App
41	423.5	24.5	364	4	US-09-542-615A-172	Sequence 172, App
42	423.5	24.5	364	4	US-09-606-421B-172	Sequence 172, App
43	423.5	24.5	364	4	US-09-614-124B-783	Sequence 783, App
44	423.5	24.5	364	4	US-09-671-325-783	Sequence 783, App
45	423.5	24.5	364	4	US-09-589-184-783	Sequence 783, App

ALIGNMENTS

RESULT 1

US-09-630-983A-20
; Sequence 20, Application US/09630983A
; Patent No. 6630330
; GENERAL INFORMATION:
; APPLICANT: Porro, Danilo
; APPLICANT: Sauer, Michael
; TITLE OF INVENTION: Ascorbic Acid Production from Yeast
; FILE REFERENCE: 2028.594000
; CURRENT APPLICATION NUMBER: US/09/630,983A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 20
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-630-983A-20

Query Match 31.6%; Score 547.5; DB 4; Length 344;
Best Local Similarity 37.4%; Pred. No. 3.1e-50;
Matches 122; Conservative 58; Mismatches 91; Indels 55; Gaps 8;

QY	7	FTLSNGVKIPGVGFGTFASEGSGKGYTAVTTALTGYRHLDCAWYILNEGVGEGIRDF	66
Db	25	FTLSNGVRIPALGLGTANPHEKLAETKQAKAAIKAGYRHIDTANAYETPTVGAIKEL	84
QY	67	LKENPSVKREDIFVCTKVMNHLHYEDVLW-----SIDSLKELGLDYVDFLVHWPAA	121
Db	85	L-EDGSIKREDLFIITKVM-----PVLWDEVDRLNESLKALGLEVYVDDLLOHWPCLCF	136
QY	122	EK-----NQGEPKGPGKVKYILKDLTENPEPTWRAMEKIY---EDRKA	163
Db	137	EKIDPKGISGLVKVPTVDDSGKTMVAADGDIY-----ETKQLEKIYLDNDHRV	186
QY	164	RSIGVSNWTIADLEKMSKFAKVMPHANQIEIHPFLNEELVQYCFSKNIMPVAYSLGSG	223
Db	187	RAIGVSNFSGIEYLERLIKEKRVKPTVQVETHPFLPQMELRKFCFMHDIILLTAYSLGSH	246
QY	224	N-----QVPTTGRVSEBNTLNIAEKGNTLAQVLIANGREGYVVLPKSSNPKIESNF	279
Db	247	GAPNLKIP-----LVKKLAEKYNTGNLLIYHIROGTIVIPRSLNPFVRISSTI	296
QY	280	KSEIELSDADFEAINAVAKGRHFRFVN	305
Db	297	EPASLTQDELQELNDFGEKYVPRFID	322

RESULT 2

US-09-347-803-6

Tue Jun 1 09:41:45 2004

```
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1473
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1473

Query Match      31.6%; Score 547.5; DB 15; Length 344;
Best Local Similarity 37.4%; Pred. No. 1.3e-47;
Matches 122; Conservative 58; Mismatches 91; Indels 55; Gaps 8;

QY 7 FTLSNGVKIPGVGFTFASGSGEYTYATTALKTGYRHLDCAWYYLNEGEVGEIRDF 66
Db 25 FSLNNGVRIPALGLGTANPHEKLAETKQAVKAAKAGYRHIDTAWAYETEPFVGEAIKEL 84

QY 67 LKENPSVKREDIFVCTKYNHNLHRYEDVLW-----SIDDSLKELGLDYVDMELVHWPIAA 121
Db 85 L-EDGSIKREDLFTTTKVV-----PVLMDVDRSLNESLKALGLEYYVDLLLOHWPLCF 136

QY 122 EK-----NGQGEPKIGPDGKYVILKDLTENPEETWAMEKIY---EDRKA 163
Db 137 EKIKDPKGISGLVKTPVDDSGKTMYAADGYL-----ETYKQLEKIYLDPNDRV 186

QY 164 RSIGVSNWTIADLEKMSKFAKMPHANOIEHPFLNEELVQYCFSKNIMPVAYSPLGSQ 223
Db 187 RAIGVSNFSIEYLERLIKECRVKPTVQVETHPHLPQOMELRKFCFMDHILLTAYSPLGSH 246

QY 224 N---QVPTTGERVSENKTLNEIAEGGNTLAQVLIAGLRRGYVVLKSSNPKRIESNF 279
Db 247 GAPNLKIP-----LVKKLAEKYVNTGNDLLISYHROGTIVIPRSLNPVRISSSI 296

QY 280 KSIELSDADFEAINAVAKGRHFRVN 305
Db 297 EFASLTQDELQELNDFGEKYVVRFD 322
```

Search completed: May 28, 2004, 14:07:19
Job time : 49 secs


```
Qy 5 KFTLSNGKIPGVGFGTFASESGKGETYATVTTALKTGYRHLDCAWYLNNEGVGEGIR 64
Db 11 KYFPLSNGEQIPALGLGTWKS--SPQVVGQAVEQALDLYRHLDCAIYNEAIGATLA 68
Qy 65 DFLKENSVKREDIFVCTKWNHHRVEDVLSIDSLKRLGLDYDMFLVHWPAAEKV 124
Db 69 NAFTKG-VVKEELWITSKLSNAHDPDAVPALEKTLQDLGLDLYLHWPVVI-- 124
Qy 125 GQGPCKIG-PDGKVIKLDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKYSKPA 153
Db 125 ---QPDVGFESGQDLPPFPASLEGTWQALEKRAVDLGLCHIGVSNFSLKLEWLSMA 151
Qy 184 KVMPHANOIEHPPLNEELVQYCFKSNMPVAYSPILGSONQVPTTGERVSENKTL---- 239
Db 182 RIPPANQVELHPYLOQSDLLTFANSQMLLTAYSPILGSDR-PAAPQQAEPKLLTDPV 240
Qy 240 -NETAERKGGNTLAQVLIANGLRGYYVLPKSSNPKRIESFKSIELSDADFEAINAVAKG 298
Db 241 INGLAAGGSAQVLLAWAIQRTVTTPKSVNPERLEQNLRAADITLTDSMAKIALLD 300
Qy 299 RHREYN 305
Db 301 RHYRYS 307

RESULT 13
US-10-369-493-2292
; Sequence 2292, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2292
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2292

Query Match 31.7%; Score 548.5; DB 15; Length 315;
Best Local Similarity 40.5%; Pred. No. 8.9e-48;
Matches 130; Conservative 52; Mismatches 102; Indels 37; Gaps 8;

Qy 7 FTLSNGKIPGVGFGTFASESGKGETYATVTTALKTGYRHLDCAWYLNNEGVGEGIRDP 66
Db 12 FTLENGKIPISIGLGTWRS--GKDETNAVCAALKAGYRHIDTAHYGNEKEIGEGIRE- 68
Qy 67 LKENPSVKREDIFVCTKWNHHRVEDVLSIDSLKRLGLDYDMFLVHWPAA----- 120
Db 69 ----SGVPRDTIWTSLKWCNAHRAGLVPLALEKTLQDLNLEVIDAYLIHWPALLSGPE 124
Qy 121 -AEKNGGEPKIGPDGKGVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKX 179
Db 125 ELPNEKGE-----LIVEDVP--IEETQAMBELLETGKRVYIGISNFNNEVLDV 173
Qy 180 SKFAPVPHANOIEHPPLNEELVQYCFKSNMPVAYSPILGSONQVPTTG-ERVENKT 238
Db 174 LTKAVPTTHQELHYPLOQTEVLEHKKLQIHVSAYPLANQDAYNSDLSKLEHKT 233
Qy 239 LNEIAEKGNN--TLAQVLIANGLRGYYVLPKSSNPKRIESFKSIELSDADFEAINAVA 296
Db 234 LVDIANARGEGITPANIAISWAKRGTSVLPKSVNBSRIVSNFLYIPLTDKEMALNNIG 293
```

```
Qy 297 KGRHF-----RFVNMKD 308
Db 294 VVRRFSHGKPAKPMFVGLQD 314

RESULT 14
US-10-424-599-282510
; Sequence 282510, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 282510
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97128C.1.pap
US-10-424-599-282510

Query Match 31.6%; Score 547.5; DB 12; Length 313;
Best Local Similarity 37.2%; Pred. No. 1.1e-47;
Matches 121; Conservative 66; Mismatches 109; Indels 29; Gaps 8;

Qy 5 KFTLSNGKIPGVGFGTFASESGKGETYATVTTALKTGYRHLDCAWYLNNEGVGEGIR 64
Db 6 KFFELNLTAKIPSLGLGTWQAE--PGVVAEALTTAIQVGYRHIDCASAYKQQAIGSALK 63
Qy 65 DFLKENSVKREDIFVCTKWNHHRVEDVLSIDSLKRLGLDYDMFLVHWPAAEKV 124
Db 64 K-LFDGQVVKREDLWITSKJWCSDHAPDVKALDKLQELQDLQDLVLIHWPVPMK-- 120
Qy 125 GQGPCKIGPDGKGVILKDLTENPE--PTWAMEKIYEDRKARSIGVSNWTIADLEKMSKF 182
Db 121 -----SGTFGLNKEYLQPDIPNTWAMEALYDSKARAIQVSNFSSKKLQDPLDI 171
Qy 183 AKVMPHANOIEHPPLNEELVQYCFKSNMPVAYSPILGSONQVPTTGERVSENKTLNEI 242
Db 172 ARVLPVAVNQVELHPGWOQPKLRAFCEKIEHLSGYSPILGSPAALKSD---ILKHPVVTET 228
Qy 243 AEKGNLTIAQVLIANGLRGYYVLPKSSNPKRIESFKSIELSDADFEAINAVAKGRHFR 302
Db 229 AERLGKTPAQVALRWGLQAGHSVLPKSTNESRIKGNDFIDFWSIPQ-DLMTKISEIKQER 287
Qy 303 FVN-----MKDTFG-----YDWPDEE 318
Db 288 LVKASFFVHETYGAYRSIEDFDWGE 312

RESULT 15
US-10-369-493-1473
; Sequence 1473, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
```


; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae Gcy1 redu

; OTHER INFORMATION: ctase, having a glycine inserted at position 2 in the amino acid
; OTHER INFORMATION: sequence
US-09-734-237B-44

Query Match 33.2%; Score 574.5; DB 12; Length 313;
Best Local Similarity 40.8%; Pred. No. 1.8e-50;
Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KFTLSNGVKIPGVGFGTFASEGSKGETYATVTTALKTGVRHLDCAWYVNEGEVGEIR 64
DB 11 KILSNTGAQIPQIGLTWQSK--ENDAYKAVLTALKOGYRHDITAAIYNEDQVQGAIK 68
QY 65 DFLKENSVKREDIFVCTKWNHLYRYEDVLSDDSLKRLGLDYVDMFLVHWPPIAAEK 124
DB 69 D-----SGVPRKDIWVTSKLCWNAHAPEAVPKALEKTLKLDYLDYLIHFPVSF-KTGE 127
QY 125 GQGPKEGPGDKYVILKDLTENP-----PTWEAKEKIYEDRKARSIGV 168
DB 119 -----DPAVQMEHLHPPELQTEFVEKHKLGIHVAIYSPFGNQNTIYESKIP---KLIHEHIIQ 239
QY 169 SNWTIADLEK--SKFAKVMFPHANQIEIHPFLPNEELVQYCFKSNIMPVAYSPILGSONQV 226
DB 170 SNFSINNKLKLLASQGNKLTAAQVVEIHPPLPQDELINCKSGIIVVEAYSPILGS--- 225
QY 227 PTTGERSVSENKTLNEIAEKGNTLAQVLIAGLRRGVVLPKSNPKRIESNFKSIELSD 286
DB 226 --TDAPLKEVPIELIAKNNQVPGHVVISWHVQGVVLPKSNVDPRIKTNKRIFTLST 283
QY 287 ADFAINAVAK 297
DB 284 EDFAINNISK 294

RESULT 8

US-10-369-493-2285
; Sequence 2285, Application US/10369493
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 2285

; LENGTH: 321

; TYPE: PRT

; ORGANISM: Schizosaccharomyces pombe

US-10-369-493-2285

Query Match 32.8%; Score 568; DB 15; Length 321;

Best Local Similarity 43.0%; Pred. No. 8.8e-50;

Matches 131; Conservative 59; Mismatches 89; Indels 26; Gaps 10;

QY 7 FTLSNGVKIPGVGFGTFASEGSKGETYATVTTALKTGVRHLDCAWYVNEGEVGEIRDF 66

DB 16 FTLADGSKIPGLGLTWRSE--PNQTKNAVKTALQGYRHDITAAAIYGNEDVGDGIKE- 72
QY 67 LKENSVKREDIFVCTKWNHLYRYEDVLSDDSLKRLGLDYVDMFLVHWPPIAAEKNGQ 126
DB 73 ----SGVPRKDIWVTSKLCWNAHAPEAVPKALEKTLKLDYLDYLIHFPVSF-KTGE 127
QY 127 GEPKIGPGDKYVILKDLTENP-EPTWRAVEKIYEDRKARSIGVSNWTIADLEKMSKFAKV 185
DB 128 DRFPKDKGNLIYER---NPIETWKAKEKLTGKVRHIGLSNFDNTNLERILKAVK 183
QY 186 MPAHQIEIHPFLPNEELVQYCFKSNIMPVAYSPILGSON---CVPTTGERVSENKTLN 240
DB 184 KPAVQMEHLHPPELQTEFVEKHKLGIHVAIYSPFGNQNTIYESKIP---KLIHEHIIQ 239
QY 241 EIAE-KG-GNTLAQVLIAGLRRGVVLPKSNPKRIESNFKSIELSDADFAINAVAK 298
DB 240 KIAKSGEGVGTATIAVSWAITRGTSVIPKSVNEQRIKSNPKYIPLTKEDMDINSI--G 297
QY 299 RHFRF 303
DB 298 IRARF 302

RESULT 9

US-09-801-368-438
; Sequence 438, Application US/09801368
; Patent No. US20020128250A1

; GENERAL INFORMATION:

; APPLICANT: Busby, Robert

; APPLICANT: Cali, Brian

; APPLICANT: Hecht, Peter

; APPLICANT: Holtzman, Doug

; APPLICANT: Madden, Kevin

; APPLICANT: Maxon, Mary

; APPLICANT: Milne, Todd

; APPLICANT: No. US20020128250A1man, Thea

; APPLICANT: Royer, John

; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir

; APPLICANT: Silva, Jeff

; APPLICANT: Summers, Eric

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

; FILE REFERENCE: 109272.147

; CURRENT APPLICATION NUMBER: US/09/801,368

; CURRENT FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 09/487,558

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: US 60/160,587

; PRIOR FILING DATE: 1999-10-20

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 438

; LENGTH: 312

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-801-368-438

Query Match 32.1%; Score 555; DB 9; Length 312;

Best Local Similarity 40.1%; Pred. No. 1.9e-48;

Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;

QY 2 SNGKFTLSNGVKIPGVGFGTFASEGSKGETYATVTTALKTGVRHLDCAWYVNEGEVGE 61

DB 7 NSSATLKLTNGASIPVLGFGTWSVDNNG--YHVSIAALKAGYRHDITAAAIYLNEEEEVGR 64

QY 62 GIRDFLKENPVSFKREDIEVCTKWNHLYRYEDVLSDDSLKRLGLDYVDMFLVHWPPIAA 121

DB 65 AIKD-----SGVPRKDIWVTSKLCWNAHAPEAVPKALEKTLKLDYLDYLIHFPVPL 117

QY 122 EKNKGGEKIPGDKYVILKDLTE-----NPEETWRAVEKIYEDRKARSIGVSNW 171

DB 118 KTRV-----TDGNVLGIPTLEDGTIDTKEWNFIKTWELMOELPRTGKTGKAVGSNF 171

QY 227 PTTGERSVSEKTLNEIAEKGNTLAQVLIWGLRGVYVLPKSNPKRIEISNFKSIELSD 286
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3396

Query Match 56.2%; Score 972.5; DB 15; Length 255;
Best Local Similarity 68.2%; Pred. No. 8.9e-92;
Matches 174; Conservative 35; Mismatches 45; Indels 1; Gaps 1;

QY 16 PGVGFGEFASGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGEVGGIRDFLKENSVKR 75
Db 1 PANGFGTFANAGAKGETYAAVKALEVGYRHLDCAWYLYNEGEVGGIRDFLKENSVKR 60
QY 76 EDIFVCTKWNHHRHRYEDVLSIDSLKRLGLDYVDMFLVHWPFAEKN-GQEPKIGPD 134
Db 61 EDIFCTKWNHHRHRYEDVLSIDSLKRLGLDYVDMFLVHWPFAEKN-GQEPKIGPD 120
QY 135 GKVVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKSKFAKVMPHANQIEI 194
Db 121 GYIIRKELTENPEPTWRAMEDVDAKTRISIGVSNWTIADLEKSKFAKVMPHANQIEI 180
QY 195 HPFLPNEELVQYCFKSNIMPVAYSPGSGNQVPTTGERVSENKTLNEIAEKGNTLAQVL 254
Db 181 HPFLPNEELVQYCFKSNIMPVAYSPGSGNQVPTTGERVSENKTLNEIAEKGNTLAQVL 240
QY 255 IAWGLRGYVVLPS 269
Db 241 LAWGLRGYVVLPS 255

RESULT 5
US-09-734-237B-42
; Sequence 42, Application US/09734237B
; Publication No. US20030064432A1
; GENERAL INFORMATION:
; APPLICANT: Rozzelli, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B593:40608
; CURRENT APPLICATION NUMBER: US/09734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-42

Query Match 33.2%; Score 574.5; DB 12; Length 312;
Best Local Similarity 40.8%; Pred. No. 1.8e-50;
Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFTLSNGVKIPGVGFTFASGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGEVGGIR 64
Db 10 KILSLNTGAQIPQIGLGTWQSK--ENDAYKAVLTALKDGYRHIDTAIYRNEDQVGGAIK 67
QY 65 DFLKENSVKREDIFVCTKWNHHRHRYEDVLSIDSLKRLGLDYVDMFLVHWPFAEKN 124
Db 68 D-----SGVPREEIFVTTKLMCTQHEPEV--ALDQSLKRLGLDYVDMFLVHWPFAEKN 117
QY 125 GQGEPKIGPDGKYVILKDLTENPE-----PTWRAMEKIYEDRKARSIGV 168
Db 118 -----DPAYIKNEDILSVPTKDGSRVDTNNWFIKTWELMQELPKTKTKAVGV 168
QY 169 SNWTIADLEKM--SKFAKVMPHANQIEIHPFLPNEELVQYCFKSNIMPVAYSPGSGNQV 226
Db 169 SNFSINNKLKLLASQGNKLTTPAANQVETHPILLPODELINFCCKSGIIVVEAYSPLGS 224
QY 227 PTTGERSVSEKTLNEIAEKGNTLAQVLIWGLRGVYVLPKSNPKRIEISNFKSIELSD 286
Db 225 --TDAPLKEPVEIIEIAKKNVQPGHVVISHWQVGVYVLPKSNPNDAIKTKRIFTLST 282

RESULT 7
US-09-734-237B-44
; Sequence 44, Application US/09734237B
; Publication No. US20030064432A1
; GENERAL INFORMATION:
; APPLICANT: Rozzelli, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B593:40608
; CURRENT APPLICATION NUMBER: US/09734,237B

QY 227 PTTGERSVSEKTLNEIAEKGNTLAQVLIWGLRGVYVLPKSNPKRIEISNFKSIELSD 286
Db 225 --TDAPLKEPVEIIEIAKKNVQPGHVVISHWQVGVYVLPKSNPNDAIKTKRIFTLST 282
QY 287 ADFAINAVAK 297
Db 283 EDFAINNISK 293

RESULT 6
US-10-369-493-22372
; Sequence 22372, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22372
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22372

Query Match 33.2%; Score 574.5; DB 15; Length 312;
Best Local Similarity 40.8%; Pred. No. 1.8e-50;
Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFTLSNGVKIPGVGFTFASGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGEVGGIR 64
Db 10 KILSLNTGAQIPQIGLGTWQSK--ENDAYKAVLTALKDGYRHIDTAIYRNEDQVGGAIK 67
QY 65 DFLKENSVKREDIFVCTKWNHHRHRYEDVLSIDSLKRLGLDYVDMFLVHWPFAEKN 124
Db 68 D-----SGVPREEIFVTTKLMCTQHEPEV--ALDQSLKRLGLDYVDMFLVHWPFAEKN 117
QY 125 GQGEPKIGPDGKYVILKDLTENPE-----PTWRAMEKIYEDRKARSIGV 168
Db 118 -----DPAYIKNEDILSVPTKDGSRVDTNNWFIKTWELMQELPKTKTKAVGV 168
QY 169 SNWTIADLEKM--SKFAKVMPHANQIEIHPFLPNEELVQYCFKSNIMPVAYSPGSGNQV 226
Db 169 SNFSINNKLKLLASQGNKLTTPAANQVETHPILLPODELINFCCKSGIIVVEAYSPLGS 224
QY 227 PTTGERSVSEKTLNEIAEKGNTLAQVLIWGLRGVYVLPKSNPKRIEISNFKSIELSD 286
Db 225 --TDAPLKEPVEIIEIAKKNVQPGHVVISHWQVGVYVLPKSNPNDAIKTKRIFTLST 282

RESULT 7
US-09-734-237B-44
; Sequence 44, Application US/09734237B
; Publication No. US20030064432A1
; GENERAL INFORMATION:
; APPLICANT: Rozzelli, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B593:40608
; CURRENT APPLICATION NUMBER: US/09734,237B

QY 61 EGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLSIDSLKRLGLDYDMFLVHPPIA 120
DB 61 EGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLSIDSLKRLGLDYDMFLVHPPIA 120
QY 121 AEKNGQEPKIPGDKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
DB 121 AEKNGQEPKIPGDKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
QY 181 KFAKVMPHANQIEIHPFLPNEELVOYCFSKNIMPVAYSPGSONOVPTTGERVSENKTLN 240
DB 181 KFAKVMPHANQIEIHPFLPNEELVOYCFSKNIMPVAYSPGSONOVPTTGERVSENKTLN 240
QY 241 EIAKGGNTLAQVLIANGLRGYVLPKSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
DB 241 EIAKGGNTLAQVLIANGLRGYVLPKSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
QY 301 FRFVNKMTFGYDVMPEETAKNLSA 325
DB 301 FRFVNKMTFGYDVMPEETAKNLSA 325

RESULT 2

US-10-327-108-3

; Sequence 3, Application US/10327108

; Publication No. US20030186400A1

; GENERAL INFORMATION:

; APPLICANT: ASAKO, HIROYUKI

; APPLICANT: WAKITA, RYUHEI

; APPLICANT: ITOH, NOBUYA

; TITLE OF INVENTION: METHOD FOR PRODUCING OPTICALLY ACTIVE

; FILE REFERENCE: 073517

; CURRENT APPLICATION NUMBER: US/10/327,108

; CURRENT FILING DATE: 2002-12-24

; PRIOR APPLICATION NUMBER: JP 2001-395884

; PRIOR FILING DATE: 2001-12-27

; PRIOR APPLICATION NUMBER: JP 2001-395885

; PRIOR FILING DATE: 2001-12-27

; PRIOR APPLICATION NUMBER: JP 2002-107648

; PRIOR FILING DATE: 2002-04-10

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 3

; TYPE: PRT

; ORGANISM: Penicillium citrinum

US-10-327-108-3

Query Match 100.0%; Score 1731; DB 14; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.5e-170;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNGKTTLSNGVKIPGVGTFASEGSKGETYTAVTALTGTGRHLDCAWYLYNEGEVG 60
DB 1 MSNGKTTLSNGVKIPGVGTFASEGSKGETYTAVTALTGTGRHLDCAWYLYNEGEVG 60

QY 61 EGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLSIDSLKRLGLDYDMFLVHPPIA 120
DB 61 EGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLSIDSLKRLGLDYDMFLVHPPIA 120

QY 121 AEKNGQEPKIPGDKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
DB 121 AEKNGQEPKIPGDKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180

QY 181 KFAKVMPHANQIEIHPFLPNEELVOYCFSKNIMPVAYSPGSONOVPTTGERVSENKTLN 240
DB 181 KFAKVMPHANQIEIHPFLPNEELVOYCFSKNIMPVAYSPGSONOVPTTGERVSENKTLN 240

QY 241 EIAKGGNTLAQVLIANGLRGYVLPKSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
DB 241 EIAKGGNTLAQVLIANGLRGYVLPKSNPKRIESNFKSIELSDADFEAINAVAKGRH 300

QY 301 FRFVNKMTFGYDVMPEETAKNLSA 325
DB 301 FRFVNKMTFGYDVMPEETAKNLSA 325

RESULT 3

US-10-608-533-1

; Sequence 1, Application US/10608533

; Publication No. US20040091902A1

; GENERAL INFORMATION:

; APPLICANT: ASAKO, HIROYUKI

; APPLICANT: SHIMIZU, MASATOSHI

; TITLE OF INVENTION: MODIFIED REDUCTASE AND ITS GENE

; FILE REFERENCE: Q76265

; CURRENT APPLICATION NUMBER: US/10/608,533

; CURRENT FILING DATE: 2003-06-30

; PRIOR APPLICATION NUMBER: JP 2002-193074

; PRIOR FILING DATE: 2002-07-02

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 1

; LENGTH: 325

; TYPE: PRT

; ORGANISM: Penicillium citrinum

US-10-608-533-1

Query Match 100.0%; Score 1731; DB 16; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.5e-170;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNGKTTLSNGVKIPGVGTFASEGSKGETYTAVTALTGTGRHLDCAWYLYNEGEVG 60
DB 1 MSNGKTTLSNGVKIPGVGTFASEGSKGETYTAVTALTGTGRHLDCAWYLYNEGEVG 60

QY 61 EGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLSIDSLKRLGLDYDMFLVHPPIA 120
DB 61 EGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLSIDSLKRLGLDYDMFLVHPPIA 120

QY 121 AEKNGQEPKIPGDKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
DB 121 AEKNGQEPKIPGDKVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180

QY 181 KFAKVMPHANQIEIHPFLPNEELVOYCFSKNIMPVAYSPGSONOVPTTGERVSENKTLN 240
DB 181 KFAKVMPHANQIEIHPFLPNEELVOYCFSKNIMPVAYSPGSONOVPTTGERVSENKTLN 240

QY 241 EIAKGGNTLAQVLIANGLRGYVLPKSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
DB 241 EIAKGGNTLAQVLIANGLRGYVLPKSNPKRIESNFKSIELSDADFEAINAVAKGRH 300

QY 301 FRFVNKMTFGYDVMPEETAKNLSA 325
DB 301 FRFVNKMTFGYDVMPEETAKNLSA 325

RESULT 4

US-10-369-493-3396

; Sequence 3396, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10452052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 14:03:12 ; Search time 48 Seconds
(without alignments)
1890.752 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1731

Sequence: 1 MSNGKFTLLSNGVKIFGVGF.....MKDTFGYDWPBETAKNLSA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1731	100.0	325	14	US-10-004-115A-1
2	1731	100.0	325	14	Sequence 1, Appli
3	1731	100.0	325	16	Sequence 3, Appli
4	972.5	56.2	255	15	Sequence 1, Appli
5	574.5	33.2	312	12	Sequence 3396, Ap
6	574.5	33.2	312	12	Sequence 42, Appli
7	574.5	33.2	313	12	Sequence 22372, A
8	568	32.8	321	15	Sequence 44, Appli
9	555	32.1	312	9	Sequence 2285, Ap
10	555	32.1	312	15	Sequence 1683, Ap
11	555	32.1	312	12	Sequence 438, App
12	553	31.7	327	15	Sequence 39, Appli
13	548.5	31.9	315	15	Sequence 2759, Ap
14	547.5	31.6	313	12	Sequence 2292, Ap
15	547.5	31.6	344	15	Sequence 282510
					Sequence 1473, Ap

16	547	31.6	328	13	US-10-040-416-6
17	545.5	31.5	328	13	US-10-040-416-4
18	545	31.5	315	12	US-10-424-599-183977
19	543.5	31.4	313	10	US-09-882-691-6
20	543.5	31.4	313	12	US-10-424-599-221378
21	541	31.3	322	12	US-09-734-237B-32
22	538	31.1	330	9	US-09-800-487A-2
23	538	31.1	330	13	US-10-040-416-2
24	534.5	30.9	313	12	US-10-424-599-221377
25	530	30.6	277	15	US-10-369-493-16605
26	525.5	30.4	304	12	US-10-424-599-252130
27	525.5	30.4	305	12	US-10-425-114-43449
28	525	30.3	279	12	US-10-282-122A-46356
29	524.5	30.3	274	15	US-10-369-493-16462
30	523.5	30.2	313	10	US-09-882-691-25
31	523.5	30.2	341	12	US-10-425-114-48430
32	521	30.1	300	12	US-10-282-122A-52772
33	518.5	30.0	308	10	US-09-882-691-10
34	518.5	30.0	339	12	US-10-425-114-59966
35	516	29.8	279	12	US-10-282-122A-45468
36	516	29.8	280	10	US-09-769-787-49
37	516	29.8	280	12	US-10-282-122A-74062
38	516	29.8	317	12	US-10-425-114-66978
39	513	29.6	329	12	US-10-425-114-54493
40	513	29.6	379	12	US-10-425-114-59318
41	512	29.6	280	9	US-09-815-242-13290
42	511.5	29.5	368	16	US-10-038-854-127
43	509.5	29.4	309	10	US-09-882-691-14
44	509	29.4	298	15	US-10-369-493-13144
45	508.5	29.4	281	15	US-10-369-493-18317

ALIGNMENTS

RESULT 1

US-10-004-115A-1
; Sequence 1, Application US/10004115A
; Publication No. US20030134402A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: MATSUMURA, KENJI
; APPLICANT: SHIMIZU, MASATOSHI
; APPLICANT: ITO, NOBUYA
; APPLICANT: WAKITA, RUTHEI
; TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
; FILE REFERENCE: 7372-72249
; CURRENT APPLICATION NUMBER: US/10/004,115A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: JP 2000-372704
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: JP 2001-006144
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: JP 2001-026594
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: JP 2001-175175
; PRIOR FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Penicillium citrinum
US-10-004-115A-1

Query Match 100.0%; Score 1731; DB 14; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.5e-170;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNGKFTLLSNGVKIPGVGFTFASGSGKGETYTAVTTALKTGVRHLDCAWYILNEGEVG 60

Db 1 MSNGKFTLLSNGVKIPGVGFTFASGSGKGETYTAVTTALKTGVRHLDCAWYILNEGEVG 60

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003544; AAF50039.2; -
DR EMBL; BT003280; AAO25037.1; -
DR HSSP; P80276; 1AH4.
DR FlyBase; FBgn036182; CG6084.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETDTASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDOKETO REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDOKETO REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO REDUCTASE_3; 1.
SQ SEQUENCE 316 AA; 35922 MW; 54745E0B7487F48B CRC64;
Query Match 30.7%; Score 531.5; DB 5; Length 316;
Best Local Similarity 38.6%; Pred. No. 3.2e-35;
Matches 124; Conservative 60; Mismatches 108; Indels 29; Gaps 10;
QY 10 SNGVKIPGVGFTFASGSGKETVAVTTALKTYRHLDCAWYLVNEGEVGEIGRDFLKE 69
DB 9 NNGKEVPIIGLTWGS--PKQVTEAVKVAIDAGYRHIDCAYYVQNEDEVDGVEAKIKE 66
QY 70 NPSVKREDIFVCTKWNHLHRYEDVLWSIDSLKRLGLDYLVDMLVHWPIDAAENKGGGP 129
DB 67 G-VYKREDLFTSKLWTFHRPDLVKSALNTLSSLKLYLDLYLIHWPV-GYKGGCDLF 124
QY 130 KIGPDGK-----YVILKOLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMSKF 182
DB 125 PTDKDGKTLXSPVDV-----DTWKAMEKLVEGLVKSIGVSNFNRRIERVLEV 174
QY 183 AKVPHANQIETHPLENEELVQYCFKSNIMFVAYSPGSGNQV-PTTGERV-SENKTLN 240
DB 175 ATIPVNTQIBCHPYLTQKLDIFCKSDITITAYSPGSPNRPWAKAGDPVILEAKIK 234
QY 241 EIAKGGNTLAQVLIAGLRRGVVLPKSNPKRIEENFK--STELSDADFEAINAV-AK 297
DB 235 EIAKKKKTPOQLIRVQQRANIVIPKSVTKDRIENFQVDFELTPSEIIEISFECH 294
QY 298 GRHFRFVNKMDTFQYDVWPEE 318
DB 295 G---RLVPELLNQYGHPPHPE 312

RESULT 11

Q81B04 PRELIMINARY; PRT; 279 AA.
AC Q81B04;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE 2,5-diketo-D-gluconic acid reductase (EC 1.1.1.-).
GN EC3392.
OS *Bacillus cereus* (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapral V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haselkorn R., Forstein M., Ehrlich S.D.,
RA Overbeek R., Kyrtides N.;
RT "Genome sequence of *Bacillus cereus* and comparative analysis with
RT *Bacillus anthracis*,"
RL Nature 423:87-91(2003).
DR EMBL; AE017009; AAP10328.1; -
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo ket red; 1.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDOKETO REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDOKETO REDUCTASE_2; 1.

DR PROSITE; PS00063; ALDOKETO REDUCTASE_3; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 279 AA; 31836 MW; D34C584585BF85CF CRC64;

Query Match 30.6%; Score 530.5; DB 16; Length 279;
Best Local Similarity 40.3%; Pred. No. 3.2e-35;
Matches 120; Conservative 56; Mismatches 85; Indels 37; Gaps 9;

QY 1 MSNGKFTT-LSNGVKIPGVGFTFASGSGKETVAVTTALKTYRHLDCAWYLVNEGEV 59
DB 1 MNLQSKTVLSNGVEMPWFGLGVFKVEEGP-ELVEAKSAIKAGYRSIDTAAIYGNAAV 59
QY 60 GEGIRDFLENPSVKREDIFVCTKWNHLHRYEDVLWSIDSLKRLGLDYLVDMLVHWPID 119
DB 60 GEGIRAGI-ETTGSRBELFITSKVNADQGVETIAAYEESLKKLKDYLGLYVHWPV 118
QY 120 AAENKQGEPEKIPGDPKGVYVLPKSNPKRIEENFK--STELSDADFEAINAV 179
DB 119 -----EGKY---KD-----TWRALETLYKEKRVFAIGVSNFQIHHLQDV 154
QY 180 SKFAKMPHANOIETHPLENEELVQYCFKSNIMFVAYSPGSGNQVPTTGERVSENKTL 239
DB 155 IDAEIKKPMINOVEYHPLRTQKELQAFCKEQIQMEAWSPL-MQGG-----LIDNETL 206
QY 240 NEIAKGGNTLAQVLIAGLRRGVVLPKSNPKRIEENFK--STELSDADFEAINAV 295
DB 207 QSIABKHGKTTAQVILRWDLQNGVITIPKSTKEHRIIANADVNFELTKEDMEKIDAL 264
RESULT 12
Q9HGX9 PRELIMINARY; PRT; 310 AA.
AC Q9HGX9;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE GlyceroL dehydrogenase.
GN ZGCY2.
OS *Zygosaccharomyces rouxii* (Candida mogii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Zygosaccharomycetes.
OX NCBI_TaxID=4956;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC42981.
RA Iwaki T., Kuroko S., Yokose Y., Kubota K., Tamai Y., Watanabe Y.;
RT "Cloning of glycerol-3-phosphate dehydrogenase genes (ZrGPD1 and
RT ZrGPD2) and glycerol dehydrogenase genes (ZrGCV1 and ZrGCV2) from the
RT salt-tolerant yeast *Zygosaccharomyces rouxii*,"
RL Yeast 010-0(2001).
DR EMBL; AB047397; BAB11960.2; -
DR HSSP; P06632; 1HW6.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETDTASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDOKETO REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDOKETO REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO REDUCTASE_3; 1.
SQ SEQUENCE 310 AA; 34106 MW; 8D1BD1424775C375 CRC64;

Query Match 30.6%; Score 530; DB 3; Length 310;
Best Local Similarity 38.7%; Pred. No. 4.2e-35;
Matches 120; Conservative 53; Mismatches 95; Indels 42; Gaps 8;

QY 6 TPTLSNGVKIPGVGFTFASGSGKETVAVTTALKTYRHLDCAWYLVNEGEVGEIGRD 65
DB 10 TLKLTGTQTIPOVGLGTWRSENEG--YKAVIEALKAGYRHIDGAAYVNEGEVGAIQD 67
QY 66 FLKENPSVKREDIFVCTKWNHLHRYEDVLWSIDSLKRLGLDYLVDMLVHWPID 120
DB 68 -----SGVPRNEIFLTKWCTQQRNPQ--ALDQSLQRLGLDYLGLYVHWPVLRTE 120

QY 7 FTLSNGVKIPGVGFGTFASEGSGKGETYAVTTTALKTGYRHLDCAWYVYLNAGEVGEGRDF 66
DB 15 FTLENGKXISIGLGTWRS--GDETKNAVCAALKAGYRHIDTAHYGNEKEIGEGIRE- 71
QY 67 LKENPSVKREDIFVCTKVNHLHRYEDVLMSIDSLKRLGLDYDMLVHPWTA----- 120
DB 72 ----SGVPRTDIWTSLKWCNAHRAGLVPLALEKTLQDLNLEVIDAYLIHWP FALLSGPE 127
QY 121 -AEKNGGPEKIPGDGVVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKM 179
DB 128 ELPNEKE-----LYEDVP--IETWQAMELLTGTQKRVIGISFNNEVILDRV 176
QY 180 SKFAYMPHANQIBIHPFLPNEELVQVCFCKNIMPVAYSPGSGNQVPTTG-ERVENKT 238
DB 177 LKIAKVPPTIHQMLHPYLPQTEYLEKHKKLQIHVSAYSPLANQNDAYNSDIKLEHKT 236
QY 239 LNETAEKGN--TLAQLIANGLRGVVLPKSNPKRISNPKSIELSDADFEAINAVA 296
DB 237 LVDIANARGEGITPANTAINAVAKRGTSVLPKSVNESRIVSNFLYIPLTDKMEAINNIG 296
QY 297 KGRHF-----RFVNMKD 308
DB 297 VVRFSGKPAKPMFVGLQD 317

RESULT 7

Q7ZA52 ID Q7ZA52 PRELIMINARY; PRT; 312 AA.
AC Q7ZA52;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aldose reductase.
OS Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
OX NCBI_TaxID=148305;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng Y.-J., Tao Y., Zhang W., Jordan D.;
RT "Inhibition of fungal aldose reductase."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033888; AAK55762.1; -.
SQ SEQUENCE 312 AA; 34368 MW; 10FE14BC676B5C55 CRC64;

Query Match 31.4%; Score 543; DB 3; Length 312;
Best Local Similarity 42.0%; Pred. No. 3.66-36;
Matches 137; Conservative 53; Mismatches 110; Indels 26; Gaps 12;
QY 5 KFTLSNGVKIPGVGFGTFASEGSGKGETYAVTTTALKTGYRHLDCAWYVYLNAGEVGEGR 64
DB 4 KDFTLNTGAKIPAFGLGTW--QGDKGVIKEAVLTAIKSGYRLIDGAVVYGNEEVCGGIR 61
QY 65 DFLKENSFKREDIFVCTKW-NHLHRYEDVLMSIDSLKRLGLDYDMLVHPWTAAX 123
DB 62 EAISSG-IVKREDLFVSKWATYTTTCE---LGLDQSLKLLGLDYDMLVHPWILMNP 117
QY 124 NQGE--PKIGPDCKYVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKMSK 181
DB 118 EGNDEKPK-KADSGRDIH--THNHVDTWKLMKLPATGKTRAVGVSNSKAWLQQLP 174
QY 182 FAKYNPHANQIETIHPFLPNEELVQVCFCKNIMPVAYSPGSGNQVPTTGERSVENTKNE 241
DB 175 HATTVPVAVNQVENHFPQLPQQLVDFCKEKGHIHMAVSPGSG-----TGGPILLTAEPVVK 228
QY 242 IAEKGGNTLAQVLIANGLRGVVLPKSNPKRISNPKSIELSDADFEAIN---AVAK 297
DB 229 IAEKHSISPAVLGYSQIARTIIVPKSVNPDRIKANAQLKDLDAEDMKLINDYSEOLAK 288
QY 298 -GRHFRFVNMDTGYDV-WPEETAK 321
DB 289 DGKLNRYV--PPRFGTDFGPDKVGQ 312

RESULT 8

Q8LBG6 ID Q8LBG6 PRELIMINARY; PRT; 309 AA.
AC Q8LBG6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative NADPH dependent mannose 6-phosphate reductase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troughan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 0:0-0 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Broeyer V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY087223; AAM64779.1; -.
DR InterPro; IPR001395; ALDO/ket red.
DR Pfam; PF00248; aldo ket red; I.
DR PRINTS; PR00069; ALDKETEDTASE.
DR PRODOM; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00798; ALDO-KETO-REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDO-KETO-REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDO-KETO-REDUCTASE_3; 1.
SQ SEQUENCE 309 AA; 34962 MW; 1FB3017045D86015 CRC64;
Query Match 30.8%; Score 533.5; DB 10; Length 309;
Best Local Similarity 40.1%; Pred. No. 2.1e-35;
Matches 120; Conservative 54; Mismatches 11; Indels 13; Gaps 7;
QY 8 TLSNGVKIPGVGFGTFASEGSGKGETYAVTTTALKTGYRHLDCAWYVYLNAGEVGEGRDF 67
DB 4 TLSNGFKMPIVGLGVWRME--KEGIRDILNALKIGYRHLDCAADYRNTEVGDALTEAF 61
QY 68 KENPSVKREDIFVCTKVNHLHRYEDVLMSIDSLKRLGLDYDMLVHPWTAAXKNGQ 127
DB 62 KTG-LVKREDLFTTKLWNSDGH--VIEACKDSLKQLDLYDLFLVHPVATKHTGVG 118
QY 128 --EPKIGPDCKYVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKMSKFAKV 185
DB 119 TTDSALGDDG--VLDDITTSLETTTHDMEXLVMGLVRSIGISNYDYFLTRDCLAYSKI 176
QY 186 MPHANQIEIHPFLPNEELVQVCFCKNIMPVAYSPGSGNQVPTTGERSV--ENKTLNEIA 243
DB 177 KEAVNQIETHPTFQDRLVKFCQKHGICVTAHTPLGGATANAEMFGTVCSDLPVLKDV 236
QY 244 EKGGGNTLAQVLIANGLRGVVLPKSNPKRISNPKSIELSDADFEAINAVAKGRH 300
DB 237 EKYKETVAVLRWGIQKTVIPIKTPARLEENFQVDFELSKEDNEVSKMERKYH 295

RESULT 9

Q96UH3 ID Q96UH3 PRELIMINARY; PRT; 314 AA.
AC Q96UH3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aldohyde reductase.
GN ALDI.
OS Coccidioides immitis.

```

RL Yeast 0:0-0(2001).
DR EMBL; AB047396; BABL1959.1; -.
DR HSP; P06632; IHW6.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETREDTASE.
DR ProDom; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE 1; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE 2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE 3; 1.
SQ SEQUENCE 310 AA; 34126 MW; F201FFBF8CCB9CFD CRC64;

Query Match
Best Local Similarity 40.5%; Score 555; DB 3; Length 310;
Matches 126; Conservative 48; Mismatches 93; Indels 44; Gaps 9;

QY 6 TFLNSGVKIPGVGFGTFASGSGKGTYYTAVTALKTGYHLCDAWYLNNEGVGEGIR 65
DB 10 TLKLTGQTTPQVGLGTRSGKNEG--YKAVIEALKAGYRHIDCAAVYGNEGVGKAIQD 67
QY 66 FLKENPSVKREDIFVCTKVM-NLHRYEDVLWSDSLKELGLDYDMFLVHWPFAIA--- 120
DB 68 -----SGVPRNEIFLTLKWTCTHQRNQEAL---DQSLQLGLDYGLVLYHWPVPRTE 119
QY 121 -----AENKQGQEPKIGPDGKYVILKLTENPEPTWRAMEKIYEDRKARSIGVSN 170
DB 120 NIKDGNLFQPAEK-----PDGSKDV--DLEWNFIKTWELMQKLLSGKTKAIGVSN 168
QY 171 WTADLEKM--SKFAKMPHANQIEHPFLNPELVQYCFSKNIMPVAYSPLGSGNQVPT 228
DB 169 FSVNNKDLALAAPTTKVTPAVNQVEHPLLPQBELIQYCEKIGVIEAYSPLGGEN---- 224
QY 229 TGERVSENKTLNEIAEKGNTLAQVLIAMGLRGRGVVLPKSSNPKRIESNFKSIELSDAD 288
DB 225 --APILSDPIVOELAKNGVAGHVVISWAVQKGLVTLPSKSVTPSLVGNLKVLTLSDD 282
QY 289 FEAINAVAKGR 299
DB 283 VAKVDGLLAK 293

RESULT 5
P74308
ID P74308 PRELIMINARY; PRT; 327 AA.
AC P74308;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE Aldehyde reductase.
GN SUR0942.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asanizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugitani M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90914; BAA18402.1; -.
DR PIR; S76143; S76143.
DR HSP; P14550; 2ALR.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETREDTASE.
DR ProDom; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE 1; 1.

DR PROSITE; PS00062; ALDOKETO_REDUCTASE 2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE 3; 1.
SQ SEQUENCE 325 AA; 36629 MW; 68ED211F1E4D08D3 CRC64;

Query Match
Best Local Similarity 40.5%; Score 548.5; DB 3; Length 325;
Matches 130; Conservative 52; Mismatches 102; Indels 37; Gaps 8;

DR PROSITE; PS00063; ALDOKETO_REDUCTASE 3; 1.
SQ SEQUENCE 327 AA; 36014 MW; 4B9415E089A8892D CRC64;

Query Match
Best Local Similarity 38.1%; Score 553; DB 16; Length 327;
Matches 117; Conservative 69; Mismatches 105; Indels 16; Gaps 6;

QY 5 KTFILNSGVKIPGVGFGTFASGSGKGTYYTAVTALKTGYHLCDAWYLNNEGVGEGIR 64
DB 11 KYFPLNSGEQIAPGLGLGTWS--SPQVVGQAQVEQALDGLGYHLCDAAYGNEAIGATLA 68
QY 65 DFLKENPSVKREDIFVCTKVM-NLHRYEDVLWSDSLKELGLDYDMFLVHWPFAIAEKN 124
DB 69 NAFKMG-VVKEELWITSLKWSNAHPDAVLPALEKTLQDLGLDYLDLYLHWPVVI--- 124
QY 125 QGQEPKIG-PDGKYVILKLTENPEPTWRAMEKIYEDRKARSIGVSNWTADLEKMSKFA 183
DB 125 ---QPDVGFPESGDQLLPFTFASLEGTMQALEKAVDLGLCHHIGVSNFSLKLEMLVLSMA 181
QY 184 KMPHANQIEHPFLNPELVQYCFSKNIMPVAYSPLGSGNQVPTTGERVSENKTL---- 239
DB 182 RIPFAVQVELHPYLOQSDLLTFANSQNILLTAVSPLGSGDR-PAATQQAAPKLLTDPV 240
QY 240 -NETAEKGGNTLAQVLIAMGLRGRGVVLPKSSNPKRIESNFKSIELSDADFEAINAVAKG 298
DB 241 INGIAAEQCSAAQVLLAWAIQRTGTTPKSNVPERLEQNLRADITLTDSEMAKIALLD 300
QY 299 RHREVN 305
DB 301 RHRYVS 307

RESULT 6
O42888
ID O42888 PRELIMINARY; PRT; 325 AA.
AC O42888;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Probable oxidoreductase.
GN SPBC8E4.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Connor R., Churher C.M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL01815; CAA16997.1; -.
DR PIR; T50378; T39169.
DR HSP; P14550; 2ALR.
DR GenedB Spombe; SPBC8E4.04; -.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETREDTASE.
DR ProDom; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE 1; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE 2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE 3; 1.
SQ SEQUENCE 325 AA; 36629 MW; 68ED211F1E4D08D3 CRC64;

Query Match
Best Local Similarity 40.5%; Score 548.5; DB 3; Length 325;
Matches 130; Conservative 52; Mismatches 102; Indels 37; Gaps 8;

```

```
QY 121 AEKNGQGEKIPGDKVKVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKMS 180
Db 121 SEKEDQEKIPGDKVKVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKLF 180
QY 181 KFAKVMHPANQIEIHPFLPNEELVQYCFKSNIMPVAYSPILGSONQVPTTGERVSENKTLN 240
Db 181 KFAEIKPHVHQIEIHPFLPNNELVQFCFKNDILPEAYSPILGSONQVPTTGERVSENKTLN 240
QY 241 EIAKGGNTLAQVLIANGLRGYYVLPKSNPNRIENFKNFISLSDADPEAINAVAKGRH 300
Db 241 EIAQKGGNTLAQVLIANGLRGYYVLPKSNPNRIENFKNFISLSDADPEAVNKAEGRH 300
QY 301 FRFVNMKDTGYDWPPEETAKNLSA 325
Db 301 FRFVNMKDTGYDWPPEETAKNLSA 325

RESULT 2
Q728L1
ID Q728L1 PRELIMINARY; PRT; 325 AA.
AC Q728L1
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE NADP(+)-dependent glycerol dehydrogenase (EC 1.1.1.72).
GN GLDB.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RA de Vries R.P.;
RT "Glycerol dehydrogenase, encoded by glbB is essential for
RT osmotolerance in Aspergillus nidulans.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ496025; CAA42649.1; -.
KW Oxidoreductase.
SQ SEQUENCE 325 AA; 36981 MW; BFADE1CB4DA15F14 CRC64;

Query Match 85.0%; Score 1471; DB 3; Length 325;
Best Local Similarity 82.5%; Pred. No. 4.1e-112;
Matches 268; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSNGKTTLSNGVKIPGVGFGTFASGSGKGETTAVTTALKTGYRHLDCAWYVLYNVEGVG 60
Db 1 MSSGKTFKLSNGVTIPAVGFGTFASGSGGETYRAVKKALEVGYRHLDCAWFYQNEDEVG 60
QY 61 EGIRDFLKENPSVKREDIFVCTKWNHLHRYEDVLSIDDSLKRLGLDYVDMFLVHPPIA 120
Db 61 DAVRDFLKENPSVKREDIFCTKWNHLHRLHPEDVRNSIEDSLKXKTDYVDLFLIHPPIA 120
QY 121 AEKNGQGEKIPGDKVKVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKMS 180
Db 121 AEKESQDKPIGDPGVKVIKDLTDEPKPTQWAMEKLYEDKARSIGVSNWTIEGLEKLL 180
QY 181 KFAKVMHPANQIEIHPFLPNEELVQYCFKSNIMPVAYSPILGSONQVPTTGERVSENKTLN 240
Db 181 KYAKVKPHVHQIEIHPFLPNEELIQYCKNDILPEAYSPILGSONQVPTTGERVSENKTLN 240
QY 241 EIAKGGNTLAQVLIANGLRGYYVLPKSNPNRIENFKNFISLSDADPEAINAVAKGRH 300
Db 241 EIAQKGGNTLAQVLIANGLRGYYVLPKSNPNRIENFKNFISLSDADPEAVNNAKGRH 300
QY 301 FRFVNMKDTGYDWPPEETAKNLSA 325
Db 301 FRFVNMKDTGYDWPPEETAKNLSA 325

RESULT 3
Q74646
ID Q74646 PRELIMINARY; PRT; 254 AA.
AC Q74646;
```

```
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Reductase (Fragment).
OS Gibberella zeae (Fusarium graminearum).
OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=5518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F15;
RX MEDLINE=98433864; PubMed=9762900;
RA Kimura M., Matsumoto G., Shingu Y., Yoneyama K., Yamaguchi I.;
RT "The mystery of the trichothecene 3-O-acetyltransferase gene, Analysis
RT of the region around Tril101 and characterization of its homologue from
RT Fusarium sporotrichoides.";
RL FEBS Lett. 435:163-168(1998).
DR EMBL; AB014493; BAA33773.1; -.
DR HSP; P06632; 1HW6.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; aldo_ket_red.1.
DR PRINTS; PR00069; ALDKETDTASE.
DR PRODOM; PD000288; Aldo/ket_red.1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
FT NON_TER 1
FT NON_TER 254
SQ SEQUENCE 254 AA; 28387 MW; 49347C4B3D5D597F CRC64;

Query Match 59.9%; Score 1037; DB 3; Length 254;
Best Local Similarity 72.8%; Pred. No. 9.7e-77;
Matches 185; Conservative 30; Mismatches 39; Indels 0; Gaps 0;

QY 16 PGVGFGTASEGSGKGETTAVTTALKTGYRHLDCAWYVLYNVEGVGIRDFLENPSYKR 75
Db 1 PGVGFGTANEGAKGETYKAVIAALKTGYRHLDCAWFYLYNVEGVQAVRDFLAEDNGYKR 60
QY 76 EDIFVCTKWNHLHRYEDVLSIDDSLKRLGLDYVDMFLVHPPIAAEKNGGGEKIPGDPG 135
Db 61 EDLFTCTKWNHLHPEEVKWSFENSLKNGFLDYLFLVHPPIAAEKGGDDYQPKIGDPG 120
QY 136 KVVILKDLTENPEPTWAMEKIYEDRKARSIGVSNWTIADLEKMSKFAKVMHPANQIEIH 195
Db 121 KVIKDKLTENPEPTWAMEIYASGKARAIGVSNWTIDGLKQLMAFAKVPANQIEIH 180
QY 196 PLPNEELVQYCFKSNIMPVAYSPILGSONQVPTTGERVSENKTLNEIAKGGNTLAQVLI 255
Db 181 PLPNEELVQFLENLIDILPAAYSPILGSONQVPTTGETVRSNKTLENAVDRSGNTLAQVLI 240
QY 256 AWGLRRGVVLPKS 269
Db 241 AWGLRRGTLFCPR 254

RESULT 4
Q9HGVO
ID Q9HGVO PRELIMINARY; PRT; 310 AA.
AC Q9HGVO;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycero1 dehydrogenase.
GN ZGCY1.
OS Zygosaccharomyces rouxii (Candida mogii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Zygosaccharomycetes.
OX NCBI_TaxID=4956;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC42981;
RA Iwaki T., Kuroso S., Yokose Y., Kubota K., Tamai Y., Watanabe Y.;
RT "Cloning of glycerol-3-phosphate dehydrogenase genes (ZrGPD1 and
RT ZrGPD2) and glycerol dehydrogenase genes (ZrGCV1 and ZrGCV2) from the
RT salt-tolerant yeast Zygosaccharomyces rouxii.";
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 14:00:41 ; Search time 45 Seconds
(without alignments)
2278.743 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1731

Sequence: 1 MSGNGKFTLSNGVKIPGVGF.....MKDTFGYDVWPEAKNLSA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMEL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1514	87.5	325	3 Q877A2	Q877A2 aspergillus
2	1471	85.0	325	3 Q728L1	Q728L1 emericeila
3	1037	59.9	254	3 Q74646	Q74646 gibberella
4	555	32.1	310	3 Q9HGf0	Q9HGf0 zygosacchar
5	553	31.9	327	16 P74308	P74308 synecocyst
6	548.5	31.7	325	3 Q42888	Q42888 schizosacch
7	543	31.4	312	3 Q7ZA52	Q7ZA52 magnaporthe
8	533.5	30.8	309	10 Q8LHG6	Q8LHG6 arabidopsis
9	533	30.8	314	3 Q6UH3	Q6UH3 coccidioid
10	531.5	30.7	316	5 Q9VTK9	Q9VTK9 drosophila
11	530.5	30.6	279	16 Q8LB04	Q8LB04 bacillus ce
12	530	30.6	310	3 Q9HGK9	Q9HGK9 zygosacchar
13	530	30.6	322	10 Q8GXW0	Q8GXW0 arabidopsis
14	529.5	30.6	309	10 Q9SJV2	Q9SJV2 arabidopsis
15	529	30.6	290	10 Q80945	Q80945 arabidopsis
16	528	30.5	280	16 Q8E460	Q8E460 streptococc

17	528	30.5	280	16 Q8DYK4	Q8DYK4 streptococc
18	527.5	30.5	309	10 Q9SJV1	Q9SJV1 arabidopsis
19	526	30.4	315	10 Q9AW93	Q9AW93 digitalis p
20	525.5	30.4	350	5 Q8IQF8	Q8IQF8 drosophila
21	525	30.3	279	16 Q81XD1	Q81XD1 bacillus an
22	525	30.3	279	16 Q81S55	Q81S55 bacillus ce
23	525	30.3	315	10 Q9AW92	Q9AW92 digitalis p
24	523.5	30.2	313	10 Q82020	Q82020 medicago sa
25	523	30.2	286	16 Q88SL1	Q88SL1 lactobacill
26	522.5	30.2	275	16 Q818Y9	Q818Y9 bacillus ce
27	522.5	30.2	311	10 Q80944	Q80944 arabidopsis
28	521	30.1	276	5 Q9GV41	Q9GV41 trypanosoma
29	521	30.1	320	10 Q9M009	Q9M009 arabidopsis
30	519	30.0	280	16 Q8DUR6	Q8DUR6 streptococc
31	518.5	30.0	275	16 Q81MD0	Q81MD0 bacillus an
32	518.5	30.0	309	10 Q8LD92	Q8LD92 arabidopsis
33	518	29.9	274	16 Q8EST2	Q8EST2 oceanobacil
34	516.5	29.8	320	5 Q9VZK8	Q9VZK8 drosophila
35	516	29.8	279	16 Q81MX4	Q81MX4 bacillus an
36	516	29.8	280	16 Q97PW2	Q97PW2 streptococc
37	516	29.8	280	16 Q8DF64	Q8DF64 streptococc
38	515.5	29.8	304	10 Q84W94	Q84W94 arabidopsis
39	508.5	29.4	281	16 Q9CIT8	Q9CIT8 lactococcus
40	507	29.3	276	5 Q8IUI5	Q8IUI5 crithidia f
41	506.5	29.3	274	5 Q965C7	Q965C7 trypanosoma
42	503	29.1	276	16 Q32210	Q32210 bacillus su
43	503	29.1	310	10 Q9FVN7	Q9FVN7 orobanche r
44	500	28.9	280	16 Q34678	Q34678 bacillus su
45	498	28.8	274	16 Q836H3	Q836H3 enterococcu

ALIGNMENTS

RESULT 1

Q877A2
ID Q877A2 PRELIMINARY; PRT; 325 AA.
AC Q877A2;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Reductase-like protein.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI840;
RA Akao T., Akeno T., Goto K., Akita O.;
RT "cloning and nucleotide sequencing of cDNA that encodes a reductase-
like protein from Aspergillus oryzae."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB079030; BAC56099.1;
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; aldo ket red; 1.
DR PRINTS; PR00069; ALDKETREDTASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
SQ SEQUENCE 325 AA; 36932 MW; DA5ED3898D0C4BD9 CRC64;

Query Match 87.5%; Score 1514; DB 3; Length 325;
Best Local Similarity 85.8%; Pred. No. 1.2e-115;
Matches 279; Conservative 22; Mismatches 24; Indels 0; Gaps 0;

QY	1	MSGNGKFTLSNGVKIPGVGF	FTASEGSKGTYTAVTTALKTGYRHLDCAWYLN	EGEVG 60
Db	1	MSGNGKFTLSNGVKIPGVGF	FTASEGSKGTYTAVTTALKTGYRHLDCAWYLN	EGEVG 60
QY	61	EGIRDFLKENSVKREDIFVCTKVNHLHRYEDV	VLWSIDSLKRLGLDYDMFLVHWP	IA 120
Db	61	DCIHDFLKENSVKREDIFVCTKVNHLHRYEDV	VLWSIDSLKRLGLDYDMFLVHWP	IA 120

QY 291 AINAVAK-GRHFFVNMKDTFGYDWWP 311
DB 295 EI---AKLDRHLEF---NDPFTWKIP 315

Search completed: May 28, 2004, 14:04:47
Job time : 19 secs

XYL1_PACTA
ID XYL1_PACTA STANDARD; PRT; 318 AA.
AC P78736;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-dependent xylose reductase [EC 1.1.1.-] (XR).
OS Pachysolen tannophilus (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pachysolen.
NCBI_TaxID=4918;
RN X
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-2460 / ATCC 32691;
RX MEDLINE=97082506; PubMed=8923742;
RA Bolen P.L., Hayman G., Shepherd H.S.;
RT "Sequence and analysis of an aldose (xylose) reductase gene from the
RT xylose-fermenting yeast Pachysolen tannophilus.";
RL Yeast 12:1367-1375(1996).
RN [2]
RP SEQUENCE OF 1-62.
RA Bolen P.L., Bietz J.A., Detroy R.W.;
RT "Aldose reductase in the yeast Pachysolen tannophilus: purification,
RT characterization and N-terminal sequence.";
RL Biotechnol. Bioeng. Symp. 15:129-148(1985).
CC !- FUNCTION: Reduces xylose into xylitol.
CC !- PATHWAY: D-xylose degradation.
CC !- SIMILARITY: Belongs to the aldo/keto reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U40706; AAC49526.1; -
DR HSP; P14550; 2ALR.
DR InterPro: IPR001395; Aldo/ket red.
DR Pfam: PF00248; aldo_ket_red; 1.
DR PRINTS; PRO0069; ALDKETRDIASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE_2; FALSE NEG.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; FALSE_NEG.
KW Oxidoreductase; Xylose metabolism; NAD.
FT ACT_SITE 46 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
FT SEQUENCE 318 AA; 36422 MW; 891CF6461D231EC CRC64;
Query Match 28.1%; Score 486; DB 1; Length 318;
Best Local Similarity 38.2%; Pred. No. 6.2e-31;
Matches 125; Conservative 51; Mismatches 117; Indels 34; Gaps 11;
QY 7 FTLSNGVPIPGVCGTFPASEGSKSETYTAVTALKTCGRHLDCAWYVINEGEVGEIRDF 66
DB 6 YTLNKRKIPAIAGMGCWKLNAADMVYAAI---KEGYRLFDCACDYGNKEVGEGINFA 61
QY 67 LKENPSVKREDIFVCTKWNHLHRYEDVLMISIDSLKRLGLDYVDMFLVHWPIAAEKNGQ 126
DB 62 IKDG-LVKKRLFTISKLNWNNFHAKENVKKALMSLSDFNLDYFDLYLMHFPISF-KFVP 119
QY 127 GEPKIGP-----DGKVIKDLTNPPEP---TWAMEKIYEDRKARSIGVSNWTIADLEK 178
DB 120 PEEKYPPPGFYCGDGFYEDV-----PIETWRAMENLVDEGLVKSIGVSNVSGGLLED 174
QY 179 MSKFAKWPFHANOIEIHHPFLPNEELVOYCFSKNMPVAYSPLGS-----QNOVPTTGR 232
DB 175 LTKARIKTPASQIEHHFYLOQNKLVYQAOLKGIWVTGYTSNFGPLSLFLGNETAKTQP 234
QY 233 VSENKTLNEIAEKGNTLAQVLIAGWLRGRGVYVLPKSNPKRIESNF--KSIELSDADFE 290
DB 235 LYENKTIITIAAKHGKTFQVLLKWNVNGRITAIIPKSTFTNTLAVNLHVDFDLTKEDFE 294

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[9] SEQUENCE OF 243-274.
RP MEDLINE=89123393; PubMed=2492527;
RX Morjana N.A., Lyons C., Flynn T.G.;
RA "Aldehyde reductase from human psoas muscle. Affinity labeling of an
RT active site lysine by pyridoxal 5'-phosphate and pyridoxal
RT 5'-diphospho-5'-adenosine.";
RL J. Biol. Chem. 264:2912-2919(1989).
[10]
RP PARTIAL SEQUENCE, AND ACETYLATION.
RC TISSUE=Muscle.
RX MEDLINE=94103588; PubMed=8281941;
RA Jacquelin M., Potier N., Klarskov K., Reymann J.-M., Sorokine O.,
RA Kieffer S., Barth P., Andrianomanga V., Biellmann J.-F.,
RA van Dorselaer A.;
RT "Sequence of pig lens aldose reductase and electrospray mass
RT spectrometry of non-covalent and covalent complexes.";
RL Eur. J. Biochem. 218:893-903(1993).
[11]
RP MUTAGENESIS OF ASP-43; TYR-48; LYS-77 AND HIS-110.
RX MEDLINE=94064641; PubMed=8245005;
RA Tarle I., Borhani D.W., Wilson D.K., Quirocho F.A., Petrash J.M.;
RA "Probing the active site of human aldose reductase. Site-directed
RT mutagenesis of Asp-43, Tyr-48, Lys-77, and His-110.";
RL J. Biol. Chem. 268:25667-25693(1993).
[12]
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX MEDLINE=92320300; PubMed=1621098;
RA Wilson D.K., Bohren K.M., Gabbay K.H., Quirocho F.A.;
RA "An unlikely sugar substrate site in the 1.65 Å structure of the
RT human aldose reductase holoenzyme implicated in diabetic
RT complications.";
RL Science 257:81-84(1992).
[13]
RP X-RAY CRYSTALLOGRAPHY (2.27 ANGSTROMS).
RX MEDLINE=93077587; PubMed=1447221;
RA Borhani D.W., Harter T.M., Petrash J.M.;
RA "The crystal structure of the aldose reductase.NADPH binary complex.";
RL J. Biol. Chem. 267:24841-24847(1992).
[14]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=94052189; PubMed=8234324;
RA Wilson D.K., Tarle I., Petrash J.M., Quirocho F.A.;
RA "Refined 1.8-Å structure of human aldose reductase complexed with the
RT potent inhibitor zopolrestat.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9847-9851(1993).
[15]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=98070237; PubMed=9405046;
RA Harrison D.H., Bohren K.M., Petsko G.A., Ringe D., Gabbay K.H.;
RA "The alrestatin double-decker: binding of two inhibitor molecules to
RT human aldose reductase reveals a new specificity determinant.";
RL Biochemistry 36:16134-16140(1997).
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide
CC variety of carbonyl-containing compounds to their corresponding
CC alcohols with a broad range of catalytic efficiencies.
CC -!- CATALYTIC ACTIVITY: Alditol + NAD(P)(+) = aldose + NAD(P)H.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DISEASE: In diabetes and galactosemia, increased AR activity leads
CC to high levels of sorbitol and galactitol, respectively, in the
CC cells of many tissues. Accumulation of sugar alcohols has been
CC shown to cause osmotic catarracts in the lens. AR is also thought
CC to play a key role in diabetic complications of three other target
CC tissues, namely, nerve, kidney and retina.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X15414; CAA33460.1; -
CC EMBL; J04795; AAS51713.1; -
DR EMBL; J05017; AAS51714.1; -
DR EMBL; M34720; AAA35560.1; -
DR EMBL; M34721; AAA35561.1; -
DR EMBL; J05474; AAA51715.1; -
DR EMBL; M59783; AAA51712.1; -
DR EMBL; M59856; AAA51712.1; JOINED.
DR EMBL; AF032455; AAB88851.1; -
DR EMBL; BC000260; AAH00260.1; -
DR EMBL; BC005387; AAH05387.1; -
DR EMBL; BC010391; AAH10391.1; -
DR PIR; A39763; A39763.
DR PDB; 1ADS; 31-OCT-93.
DR PDB; 1AHN; 31-JAN-94.
DR PDB; 2ACQ; 31-JUL-94.
DR PDB; 2ACR; 31-JUL-94.
DR PDB; 2ACS; 31-JUL-94.
DR PDB; 2ACU; 31-JUL-94.
DR PDB; 1MAR; 20-JUL-95.
DR PDB; 1AZ1; 18-MAR-98.
DR PDB; 1AZ2; 18-MAR-98.
DR PDB; 1ZF3; 07-FEB-01.
DR PDB; 1EL3; 10-MAY-00.
DR Aarhus/Ghent-2DPAGE; 1202; IEF.
DR Genew; HGNC:381; AKR1B1.
DR MIM; 103880; -
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0004032; F:aldehyde reductase activity; TAS.
DR GO; GO:0005483; F:electron transporter activity; TAS.
DR GO; GO:0005975; F:carbohydrate metabolism; TAS.
DR InterPro; IPR001395; Aldo/ket red.
DR PRINTS; PF00248; Aldo/ket red; 1.
DR PRODOM; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00062; ALDO-KETO-REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDO-KETO-REDUCTASE_3; 1.
DR PROSITE; PS00798; ALDO-KETO-REDUCTASE_1; 1.
KW Oxidoreductase; NADP; Acetylation; 3D-structure; Polymorphism.

Query Match 28.1%; Score 486.5; DB 1; Length 315;
Best Local Similarity 38.1%; Pred. No. 5.6e-31;
Matches 117; Conservative 51; Mismatches 110; Indels 29; Gaps 8;

QY 9 LNSGVKIPGVGFTFASGSGKGTVTATTALKTGTYRHLDCAWYLYNVEGVEGIRDFLK 68
DB 6 LNSGAMPILGLGTWKS--PFGQVTEAVKVAIDVGYRHIDCAHVYQNEVGVATQEKLR 63

QY 69 ENPSVKREDIFVCTKWNHLYRYEDVLWSIDSLKRLGLDYVDMFLVHWPIAAEKNGQGE 128
DB 64 EQ-VVRELEFVSKLWCTYHEKGLVKGCQKTLSDKLDLYLHLHWPTGFK----- 116

QY 129 PKIGPDGKYVILKDLTENPEP-----TWAMEKIYEDRKARSIGVSNWTADLEKM-- 179
DB 117 -----PGKEFFELDESNNVPSDNTNILDRTWAAMELVDEGLVKAIGISNFNHLQVEMILN 171

QY 180 SKFAKVMHAKQIEIHPLELVOYCFNSKINMPVAVSLGSONQVPTTGERYS--ENK 237
DB 172 KFLGYKFAVNGIECHPYLTQELIYQCSQGI VVTAYSLPGSDPRWPAKEDFSLLEDP 231

QY 238 TLNEIAEKGKNTLAQVLIAGLIRGYVYLPKSSNPKRIESNFK--SIELSDADFEAINAV 295
DB 232 RIKAJAAKHNTTAQVLIIRFPMQRLNVLVIPSVTPIAENFKVDFELSSQDMTTL--L 289

QY 296 AKGRHFR 302
DB 290 SYNRNWR 296

RESULT 15

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U37100; AAC17469.1; --
CC EMBL; AF052577; AAC36465.1; --
CC EMBL; BC008837; AAH08837.1; --
CC EMBL; AF044961; AAC15671.1; --
CC GenBank; HGNC:382; AKR1B10.
CC MIM; 604707; --
CC HSP; P45377; 1FRB.
CC GO; GO:0004033; F:aldo-keto reductase activity; TAS.
CC InterPro; IPR001395; Aldo/ket red.
CC Pfam; PF00248; aldo ket red; 1.
CC PRINTS; PR00069; ALDKETREDASE.
CC ProDom; PD000288; Aldo/ket red; 1.
CC PROSITE; PS00798; ALDO-KETO-REDUCTASE_1; 1.
CC PROSITE; PS00062; ALDO-KETO-REDUCTASE_2; 1.
CC PROSITE; PS00063; ALDO-KETO-REDUCTASE_3; 1.
CC Oxidoreductase; NADP; Polymorphism.
CC ACT SITE 49 49 HYDROGEN-BOND DONOR (BY SIMILARITY).
CC FT ACT SITE 313 313 D -> N.
CC FT VARIANT 313 313 /FTID=VAR_013287.
CC SEQUENCE 316 AA; 36020 MW; 0C2FC0E378955A33 CRC64;

Query Match 28.1%; Score 487; DB 1; Length 316;
Best Local Similarity 36.6%; Pred. No. 5.2e-31;
Matches 118; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

Qy 6 TFT-LSNGKIPGVGFGTFSESGKGTAVTATKATGVRLHLDCAWYVNECEGEGR 64
Db 3 TFVELSTKAKPIVGLTWKS--PLGKVKAEAKVADAGVRHIDCAWYVNECEGEATQ 60
Qy 65 DFLKPNPSVKREDIFVCTKVNHLHRYEDVLSIDSLKRLGLDYDMFLVHWPIAAKN 124
Db 61 EKIQEK-AVKREDLFIYSKLVPTFFERPLVRKAPEKTLKDLKSLYLDVLIHWPOQFKSG 119
Qy 125 GQGEPK-----IGPDQKYVILKDLTENPEPTWRAMEKIYEDRKASIGVSNWTIADLEK 178
Db 120 DDLFPKDDKNAIG--GKATFL-----DAWEAMEELVDEGLVKALGVSNFSHFQIEK 169
Qy 179 M--SKFAKWPHPANQIHPFLNEELVQYCFKSNIMPVAYSPLGSONQVPTTGERVS-- 234
Db 170 LLNKPGLKYKPVTVNQVECHPYLTQELIQYCHSKGITVTAYSPLGSPDRPWAKPEDPSLL 229
Qy 235 ENKTLNIEAKGGMNTLAQVLIANGRLRGYVVLPKSNPKRIESNFKSLIELSDADFAINA 294
Db 230 EDPKIEIAAKHKHTAAQVLIIRFHQIRNVIPKSVTPARIVENIQVDFPKLSDEEMATI 289
Qy 295 VAKGRHFRFVNMKDTFGYDVMP 316
Db 290 LSFNRRWACNVLOSSHLEDYP 311

RESULT 14

ALDR_HUMAN
ID ALDR_HUMAN STANDARD; PRT; 315 AA.
AC P15121; O9BS21;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aldose reductase (EC 1.1.1.21) (AR) (Aldehyde reductase).
GN AKR1B1 OR ALDR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89255461; PubMed=2498333;
RA Bohren K.M., Bullock B., Wermuth B., Gabbay K.H.;
RT "The aldo-keto reductase superfamily. cDNAs and deduced amino acid
RN sequences of human aldehyde and aldose reductases.";
J. Biol. Chem. 264:9547-9551(1989).
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal;
RX MEDLINE=90045960; PubMed=2510130;
RA Graham A., Hedge P.J., Powell S.J., Riley J., Brown L., Gammack A.,
Carey F., Markham A.F.;
RT "Nucleotide sequence of cDNA for human aldose reductase.";
Nucleic Acids Res. 17:8368-8368(1989).
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89359274; PubMed=2504709;
RA Chung S., Lamendola J.;
RT "Cloning and sequence determination of human placental aldose
RN reductase Gene.";
J. Biol. Chem. 264:14775-14777(1989).
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=90253609; PubMed=2111143;
RA Grundmann U., Bohn H., Obermeier R., Amann E.;
RT "Cloning and prokaryotic expression of a biologically active human
RN placental aldose reductase.";
DNA Cell Biol. 9:149-157(1990).
RP SEQUENCE FROM N.A.
RX MEDLINE=90277668; PubMed=2112546;
RA Nishimura C., Matsuura Y., Kokai Y., Akera T., Carper D., Morjana N.,
Lyons C., Flynn T.G.;
RT "Cloning and expression of human aldose reductase.";
J. Biol. Chem. 265:9788-9792(1990).
RP SEQUENCE FROM N.A.
RX MEDLINE=91201333; PubMed=1901857;
RA Graham A., Brown L., Hedge P.J., Gammack A.J., Markham A.F.;
RT "Structure of the human aldose reductase gene.";
J. Biol. Chem. 266:6872-6877(1991).
RP SEQUENCE FROM N.A.
RX MEDLINE=97341182; PubMed=9195951;
RA Ko B.C.B., Ruepp B., Bohren K.M., Gabbay K.H., Chung S.S.;
RT "Identification and characterization of multiple osmotic response
RN sequences in the human aldose reductase gene.";
J. Biol. Chem. 272:16431-16437(1997).
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Eye, and Urinary bladder;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL; D32250; BAA06980.1; -.
DR EMBL; L39795; AAA62176.1; -.
DR EMBL; U29152; AAA69958.1; -.
DR EMBL; U89140; AAC13358.1; -.
DR EMBL; U89142; AAC13358.1; JOINED.
DR EMBL; U89143; AAC13358.1; JOINED.
DR EMBL; U89144; AAC13358.1; JOINED.
DR EMBL; U89145; AAC13358.1; JOINED.
DR EMBL; U89146; AAC13358.1; JOINED.
DR EMBL; U89147; AAC13358.1; JOINED.
DR EMBL; U89148; AAC13358.1; JOINED.
DR EMBL; U89149; AAC13358.1; JOINED.
DR EMBL; U93231; AAD32300.1; -.
DR EMBL; U93230; AAD32300.1; JOINED.
DR EMBL; AB016665; BAA76413.1; -.
DR EMBL; AB016662; BAA76413.1; JOINED.
DR EMBL; AB016663; BAA76413.1; JOINED.
DR EMBL; AB016664; BAA76413.1; JOINED.
DR EMBL; BC004725; AAH04725.1; -.
DR EMBL; BC021655; AAH21655.1; -.
DR EMBL; I49484; I49484.
DR HSP; P15121; 2ACQ.
DR SWISS-2DPAGE; P45376; MOUSE.
DR PMMA-2DPAGE; P45376; -.
DR MGD; MGI:1353494; Akr1b1.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETREDASE.
DR PRODOM; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00062; ALDOXETO REDUCTASE 2; 1.
DR PROSITE; PS00063; ALDOXETO REDUCTASE 3; 1.
DR PROSITE; PS00798; ALDOXETO REDUCTASE 1; 1.
DR OXIDOREDUCTASE; NADP; Acetylation; Multigene family.
DR INIT_MET 0 0 BY SIMILARITY.
DR NP_BIND 9 18 NADP (POTENTIAL).
DR ACT_SITE 48 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
DR MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
DR CONFLICT 45 45 A -> S (IN REF. 1).
DR CONFLICT 220 220 A -> G (IN REF. 1).
DR CONFLICT 280 280 V -> L (IN REF. 7).
DR SEQUENCE 315 AA; 35601 MW; 607E604PF432580E CRC64;
```

```
Query Match 28.2%; Score 487.5; DB 1; Length 315;
Best Local Similarity 38.0%; Pred. No. 4.7e-31;
Matches 116; Conservative 52; Mismatches 112; Indels 25; Gaps 8;

Qy 9 LSNQKVIQGVGFGTFASGSGKGYTAVTTALKTGYRHLDCAWYILNEGVGEGIRDFLK 68
Dy 6 LNNGTQPTIGLTWKS--PPQVTEAVKVAIDLGYRHLDCAQVYQNEKVGVALQK 63
Qy 69 ENPSVKREDIVCTKWNHRLHYEDVLWISDDSLKRLGLDYDMFLVHPVFAEKNGQGE 128
Dy 64 EQ-VVKEQDLFVSKLCTFHDKSNWGAQFKTSLDLQLDLYLHWPTEG----- 115
Qy 129 PKIGPDCKYVLKDLTENPPE-----TWAMEKIYEDRKARSGVSNWTIADLEKM-- 179
Dy 116 -KPGPD--YFPL-DASGNVPSDFTVDVTWAMEQLVGLVGTGVSNTFNPQIBRLN 171
Qy 180 SKFAKVMPHANQIBIHPFLNEELVQYCFKSNIMPVAYSPLGSONQVPTTGERVS---ENK 237
Dy 172 KPGLKYKPAVNIQIECHPYLTQKLIYCHSKGIVVTAYSPGSDPRWAKPEDPSLLEDP 231
Qy 238 TLNEIAKGGNTLAQVLIAGLRGVYVLPKSNPKRIENFKSIELSDADFAINAVAK 297
Dy 232 RIKAIAKYNKTKTAQVLRFPQIQRNLVVPKSTPTVPAENLKVDFEVSSEDMATLLSY 291
Qy 298 GRHFR 302
Dy 1:::1
```

```
Db 292 NRNRW 296
RESULT 13
AKBA_HUMAN
ID AKBA_HUMAN STANDARD; PRT; 316 AA.
AC O60218; O75890;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aldo-keto reductase family 1 member B10 (EC 1.1.1.-) (Aldose
DE reductase-like) (ARL-1) (Small intestine reductase) (SI reductase)
DE (Aldose reductase-related protein) (ARP) (HARP).
GN AKR1B10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Liver tumor;
RX MEDLINE=98234319; PubMed=9565553;
RA Cao D., Fan S.T., Chung S.S.M.;
RT "Identification and characterization of a novel human aldose
RT reductase-like gene";
RL J. Biol. Chem. 273:11429-11435 (1998).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT ASN-313.
RC TISSUE=Small intestine;
RX MEDLINE=98438733; PubMed=9765596;
RA Hyndman D.J., Flynn T.G.;
RT "Sequence and expression levels in human tissues of a new member of
RT the aldo-keto reductase family.";
RL Biochim. Biophys. Acta 1399:198-202 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 232-316 FROM N.A.
RX MEDLINE=98196779; PubMed=9537432;
RA Scuric Z., Stain S.C., Anderson W.F., Hwang J.-J.;
RT "New member of aldose reductase family proteins overexpressed in human
RT hepatocellular carcinoma.";
RL Hepatology 27:943-950 (1998).
CC -!- FUNCTION: Can efficiently reduce aliphatic and aromatic aldehydes,
CC and is less active on hexoses. May be responsible for
CC detoxification of reactive aldehydes in the digested food before
CC the nutrients are passed on to other organs.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- TISSUE SPECIFICITY: Found in many tissues. Highly expressed in
CC small intestine, colon and adrenal gland.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
```

```

DR EMBL; U13691; AAB60687.1; JOINED.
DR EMBL; U13692; AAB60687.1; JOINED.
DR EMBL; U13693; AAB60687.1; JOINED.
DR EMBL; M32818; AAA31160.1; -.
DR EMBL; U12316; AAA50833.1; -.
DR EMBL; J05048; AAA31157.1; -.
DR PIR; A34406; A34406.
DR HSP; P15121; 2ACQ.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETREDTASE.
DR ProDom; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00062; ALDOKETO REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO REDUCTASE_3; 1.
DR PROSITE; PS00798; ALDOKETO REDUCTASE_1; 1.
DR OXOREDUCTASE; NADP; Acetylation.
KW INIT MET 0 0 BY SIMILARITY.
FT NP BIND 9 18 NADP (POTENTIAL).
FT ACT SITE 48 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 315 AA; 35632 MW; 64D53E6AC0853FFB CRC64;

Query Match 28.2%; Score 488.5; DB 1; Length 315;
Best Local Similarity 38.4%; Pred. No. 3.9e-31;
Matches 118; Conservative 49; Mismatches 111; Indels 29; Gaps 8;

QY 9 LSNQVKIPGVGFGFASGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGVEGEGIRDFLK 68
DB 6 LYNQAKPILGLGWTWS--PPQVTEAVKTAIDLGYRHIDCAHVQNEVEGVALQEK 63

QY 69 ENSVKGREDIPVCKVWNHHRVEDVLSDDSLKGLDLYDVMFLVHWPFAEKNGQGE 128
DB 64 EQ-VVKEELFVSKLMCTSHDKSLVKGACQKTLNDKLDLYLIHWPTGPFHGESEYF 122

QY 129 PKIGPCKYVILKDLTENPSP-----TWRAKEKIVEDRKASIGVSNWTIADLEKM-- 179
DB 123 P-----LDAAGNVIPSTDFDTWEAGLVDGLVKSIGVSNFNLQIERILN 171

QY 180 SKFAKVMPHANQIEIHPFLPNEBELVQYCFKSNIMPVAYSLGSONVPTTGERVS--ENK 237
DB 172 KPLGKYKPAVNVQIECHPYLTOEKLQYCHSKGIWTVAYSLGSPDRFWAKPEDPSLLEDP 231

QY 238 TLNEIAKGGNTLAQVLIANGRGVYVLPKSNPKRIENFK--SIELSDADFEAINAV 295
DB 232 RIKAIADKHKKTAQVLIIRPQMORNLVVPKSVTPARIAENFQVFFELSSDMTTL--L 289

QY 296 AKGRHFR 302
DB 290 SYNENWR 296

RESULT 12
ID ALDR_MOUSE STANDARD; PRT; 315 AA.
AC P45376; O70130; Q99K9;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aldose reductase (EC 1.1.1.21) (AR) (Aldehyde reductase).
GN AKR1B1 OR ALDR1 OR ALDR01.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Kidney;
RX MEDLINE=95154325; PubMed=7851421;
RA Gui T., Tanimoto T., Kokai Y., Nishimura C.;
RT "Presence of a closely related subgroup in the aldo-ketoreductase
RL family of the mouse.";
RL Eur. J. Biochem. 227:448-453 (1995).
[2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Liver;
RA Iwata T., Carper D.;
RL Submitted (MAR-1995) to the ENBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney;
RA Dacudal S., Berger M., Pailhoux E., Tournaire C., Veyssiere G.,
RA Jean C.;
RL Submitted (JUN-1995) to the ENBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX MEDLINE=98153248; PubMed=9485485;
RA McGowan M.H., Iwata T., Carper D.A.;
RT "Characterization of the mouse aldose reductase gene and promoter in a
RT lens epithelial cell line.";
RL Mol. Vision 4:2-2 (1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX Ho H.T.B., Jenkins N.A., Copeland N.G., Gilbert D.J., Winkles J.A.,
RA Louie H.W.Y., Lee F.K., Chung S.S.M., Chung S.X.;
RT "Comparisons of genomic structures and chromosomal locations of the
RT mouse aldose reductase and aldose reductase-like genes.";
RL Eur. J. Biochem. 259:726-730 (1999).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=99160426; PubMed=10049784;
RA Li H., Nobukuni Y., Gui T., Yabe-Nishimura C.;
RT "Characterization of genomic regions directing the cell-specific
RT expression of the mouse aldose reductase gene.";
RL Biochem. Biophys. Res. Commun. 255:759-764 (1999).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=breast;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide
CC variety of carbonyl-containing compounds to their corresponding
CC alcohols with a broad range of catalytic efficiencies.
CC -!- CATALYTIC ACTIVITY: Alditol + NAD(P)(+) = aldose + NAD(P)H.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Abundant in the testis, skeletal muscle and
CC kidney.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


GN P100/11E.
OS Leishmania major.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89139503; PubMed=2918000;
RA Samaras N., Spichill T.W.;
RT "The developmentally regulated p100/11E gene of Leishmania major
shows homology to a superfamily of reductase genes.";
RL J. Biol. Chem. 264:4251-4254(1999).
CC -!- DEVELOPMENTAL STAGE: P100/11E abundance is markedly elevated in
promastigotes relative to amastigotes.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
EMBL; J04483; AAA7350.1; -
DR PIR; A32950; A32950.
DR HSSP; P06632; LHW6.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETRDTASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
DR Oxidoreductase.
FT ACT SITE 54 HYDROGEN-BOND DONOR (BY SIMILARITY).
SQ SEQUENCE 284 AA; 54 82FFEB17F7EDC6CC CRC64;
Query March 28.38; Score 489.5; DB 1; Length 284;
Best Local Similarity 36.6%; Pred. No. 2.9e-31;
Matches 115; Conservative 54; Mismatches 100; Indels 45; Gaps 10;
Qy 8 TSLNGKIPGVGFTFASGSKGE-TYVAVTTALKTGVRHLDCAWYLYNEGEVGEIRDF 66
Db 10 TSLNGKVPQGLGWQSPA--GEVTENAVNWCAGYRHIDTAALYKNEESVGAGLR-- 65
Qy 67 LKENPSVKREDIFVCTKWVHLHYEDVLSIDSLKRLGLDYVDFLHWPTAAEKNGQ 126
Db 66 ---ASGVREDVFITTKLWNTGQYESTLAAFEESKQLGVYDLYLIHWP-----R 115
Qy 127 GEPKIGPDGKYLKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTADLEKMSKFAVM 186
Db 116 GKDLSEKGYL-----DSWRAFEQLYKEKKVRAIGVNFHILHLEDVLANCTVT 166
Qy 187 PHANQIIEHPLPNEELVQVCFKSNIMPVAYSLGSONQVPTTGERVSEKNTLNEAEKG 246
Db 167 PMVQVVELHPLNQAOLRAFCDAKQIKVEAWSPLGQ-----GKLIS-NPILSAIGAKY 218
Qy 247 GNTLAQVLIANGVRGVVLPKSNPKRIESEN--FKSIELSDADFAINAVAKGRHFRFV 304
Db 219 NKTAQVILRWNIQKLIITPKSVHREIRENADIFDFELGAEDVMSIDA-----L 269
Qy 305 NMKDTFGDYVWPEE 318
Db 270 NTNSRYGPD--PDE 281
RESULT 10
ALDR PIG STANDARD; PRT; 315 AA.
AC P80276;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Aldo reductase (EC 1.1.1.21) (AR) (Aldehyde reductase).
GN AKR1B1 OR ALR2.
OS Sus scrofa (Pig).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE, MASS SPECTROMETRY, AND DISULFIDE BOND.
RX TISSUE=Lens;
RX MEDLINE=94109388; PubMed=8281941;
RA Jaquinod M., Potier N., Klarskov K., Reymann J.-M., Sorokine O.,
RA Kieffer S., Barth P., Andrianomanga V., Biellmann J.-F.,
RA van Dorselaer A.;
RT "sequence of pig lens aldose reductase and electrospray mass
spectrometry of non-covalent and covalent complexes.";
RL Eur. J. Biochem. 218:893-903(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=93263021; PubMed=8493902;
RA Kubiseki T.J., Green N.C., Flynn T.G.;
RT "Location of an essential arginine residue in the primary structure
of pig aldose reductase.";
RL Adv. Exp. Med. Biol. 328:259-265(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=92131138; PubMed=1734286;
RA Rondeau J.-M., Tete-Favier F., Podjarny A., Reymann J.-M., Barth P.,
RA Biellmann J.-F., Moras D.;
RT "Novel NADPH-binding domain revealed by the crystal structure of
aldose reductase.";
RL Nature 355:469-472(1992).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=97341224; PubMed=9195881;
RA Urzhumtsev A., Tete-Favier F., Mitschler A., Barbantou J., Barth P.,
RA Urzhumtseva L., Biellmann J.-F., Podjarny A.D., Moras D.;
RT "A 'specificity' pocket inferred from the crystal structures of the
complexes of aldose reductase with the pharmacologically important
inhibitors tolrestat and sorbinil.";
RL Structure 5:601-612(1997).
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide
variety of carbonyl-containing compounds to their corresponding
alcohols with a broad range of catalytic efficiencies.
CC -!- CATALYTIC ACTIVITY: Alditol + NAD(P)(+) = aldose + NAD(P)H.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MASS SPECTROMETRY: MW=35778; MW_ERR=3; METHOD=Electrospray.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
EMBL; L14950; AAA30989.1; -
DR EMBL; U46065; AAC48515.1; -
DR PIR; A59021; A59021.
DR PDB; 1DLA; 30-APR-94.
DR PDB; 1AH0; 15-APR-98.
DR PDB; 1AH3; 15-APR-98.
DR PDB; 1AH4; 15-APR-98.
DR PDB; 1EKO; 10-MAY-00.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETRDTASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.

```

DR PRINTS; PRO0069; ALDXTETRDASE.
DR PRODOM; PD000288; Aldo/ket red; 1.
DR PROSITE; PS00798; ALDOXETO_REDUCTASE_1; 1.
DR PROSITE; PS00062; ALDOXETO_REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; 1.
KW Oxidoreductase; NADP.
FT ACT SITE 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
SQ SEQUENCE 310 AA; 34905 MW; DECCAD94BF3E6D9F CRC64;

Query Match 29.0%; Score 501.5; DB 1; Length 310;
Best Local Similarity 34.8%; Pred. No. 3.7e-32;
Matches 110; Conservative 73; Mismatches 118; Indels 15; Gaps 8;

Qy 6 TFTLSNGVKPGVCFGTFAEGSGKETYTAVTALKTGYHLDCAWYLYNEGEVGEIRD 65
Db 3 TVTLSSGYEMPIVGLGWRLK--KDELKEVILNAIKIGYRHFCAAHYKSEADVGEALAE 60
Qy 66 FLKENPVQGEEDIFVCTKQVNHLYEDVLSDDSLKRLGLDYVDMFLVHWPAAEKNG 125
Db 61 AFKGTG-LVKREELPITTKWNSDGHG--VVEACKNSLEKLIQIDYLDLYLVHYEMPTKHNA 117
Qy 126 QGEPRK--IGPDGKVILKDTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMSKFA 183
Db 118 IGTASLIGED--KVLDDIVTISLQQTWEGNEKTVSLGLVRSGLGNYELFLTRDCLAYS 175
Qy 184 KVMFHAQIETHPLPNEELVOYCFSKNIMPVAYSPLGSGNQVPTTCERVS--ENKTLNE 241
Db 176 KIKFAVSQFTHPYFQDRLSVKPCWKHGVLPTAHTPLGGAANKDMFGSVPLDDPVLND 235
Qy 242 IAEKGNLTALQVLIANGRLRGYVVLPKSNPKRIESNFKSIE--LSDADFEAINAVAKGR 299
Db 236 VAKYGYKSAQICLRWGIQRTAVIPKSSKIQRLENLEVLQSLDEDMQLIYSI--DR 293
Qy 300 HRFVNMKDTFGYDVW 315
Db 294 KYRSLFSKWTGLDLY 309

RESULT 8
ALDR RAT STANDARD; PRT; 315 AA.
AC P07943;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aldose reductase (EC 1.1.1.21) (AR) (Aldohyde reductase).
GN AKR1B1 OR ALDR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Lens;
RX MEDLINE=87276556; PubMed=3111886;
RA Carper D., Nishimura C., Shiohara T., Dietzchold B., Wistow G.,
RA Craft C., Kador P., Kinoshita J.H.;
RT "Aldose reductase and p-cristallin belong to the same protein
RL superfamily as aldohyde reductase.";
RL FEBS Lett. 220:209-213(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92084118; PubMed=1748296;
RA Graham C.E., Szpirer C., Levan G., Carper D.;
RT "Characterization of the aldose reductase-encoding gene family in
RL rat.";
RL Gene 107:259-267(1991).
RN [3]
RP SEQUENCE OF 155-168 AND 204-209.
RC TISSUE=Astrocytes;
RX MEDLINE=96007849; PubMed=7498172;
RA Laeng P., Bouillon P., Raupenot L., Labourdette G.;
RT "Long-term induction of an aldose reductase protein by basic

```

```

RT fibroblast growth factor in rat astrocytes in vitro.";
RL Electrophoresis 16:1240-1250(1995).
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide
CC variety of carbonyl-containing compounds to their corresponding
CC alcohols with a broad range of catalytic efficiencies.
CC -!- CATALYTIC ACTIVITY: Alditol + NAD(P)(+) = aldose + NAD(P)H.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X05894; CAA29308.1; -.
CC EMBL; M60322; AAA40721.1; -.
CC PIR; A60603; A60603.
CC HSSP; PLS121; 2AQC.
CC InterPro; IPR001395; Aldo/ket_red.
CC Pfam; PF00248; aldo_ket_red; 1.
CC PRINTS; PRO0069; ALDXTETRDASE.
CC PRODOM; PD000288; Aldo/ket red; 1.
CC PROSITE; PS00062; ALDOXETO_REDUCTASE_2; 1.
CC PROSITE; PS00063; ALDOXETO_REDUCTASE_3; 1.
CC PROSITE; PS00798; ALDOXETO_REDUCTASE_1; 1.
KW Oxidoreductase; NADP; Acetylation.
FT INIT MET 0 BY SIMILARITY.
FT NP BIND 9 18 NADP (POTENTIAL).
FT ACT SITE 48 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 315 AA; 35666 MW; 572941A154BC1202 CRC64;

Query Match 28.9%; Score 499.5; DB 1; Length 315;
Best Local Similarity 39.4%; Pred. No. 5.4e-32;
Matches 121; Conservative 51; Mismatches 106; Indels 29; Gaps 10;

Qy 9 LNSGVKIPGVGFTFASGSGKETVTAVTALKTGYRHLDCAWYLYNEGEVGEIRDFLK 68
Db 6 LNSGVKIPGLTGLTWKS--PFGQVTEAVKVAIDGYRHIDCAQYQNEKEVGVALQEKLK 63
Qy 69 ENPSVKREDIFVCTKQVNHLYEDVLSDDSLKRLGLDYVDMFLVHWPAAEKNGOGE 128
Db 64 EQ-VVVRQDLFTVSKLWCTFHQDSMWKGCQKTLSDQLQDLYLDLYLHWPTGF----- 115
Qy 129 PKIGPDGKVILKDLTENPEP-----TWAMEKIYEDRKARSIGVSNWTIADLEKM-- 179
Db 116 -KPGPD--YFPL-DASGNVPSDDTDFVTWTAMEQLVDEGLVKAIGVSNFNPQLIERLIN 171
Qy 180 SKFAKVMFHAQIETHPLPNEELVOYCFSKNIMPVAYSPLGSGNQVPTTCERVS--BNK 237
Db 172 KGLKVKPAVNGIECHPVLTKETKLEYCHKGIVVTAYSPILGSPDRPWAKPEDPSLLEDP 231
Qy 238 TLNEIAEKGNTLAQVLIANGRLRGYVVLPKSNPKRIESNEFK--SIELSDADFEAINAV 295
Db 232 RIKEIAAKYNTKTAQVLIIRFPIQRNLVWIPKSVTPARIAENPKVDFELSDMATL--L 289
Qy 296 AKGRHFR 302
Db 290 SYNRNWR 296

RESULT 9
P100_LEIMA
ID P100 LEIMA STANDARD; PRT; 284 AA.
AC P22045;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable reductase (EC 1.1.-.-).

```

```

RESULT 6
ID ALD1_MOUSE STANDARD; PRT; 315 AA.
AC P21300;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aldose reductase-related protein 1 (EC 1.1.1.21) (AR) (Aldehyde
DE reductase) (VAS deferens androgen-dependent protein) (MVDP)
DE (Aldo-keto reductase family 1 member B7).
DE AKR1B7 OR AVDP.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Vas deferens;
RX MEDLINE=92345100; PubMed=1637719;
RA Pailhoux E.A., Veyssi re G.M., Fabre S., Tournaire C., Jean C.G.;
RT "The genomic organization and DNA sequence of the mouse vas deferens
RT androgen regulated protein gene.";
RN J. Steroid Biochem. Mol. Biol. 42:561-568(1992).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Vas deferens;
RX MEDLINE=91060614; PubMed=2121194;
RA Pailhoux E.A., Martinez A., Veyssi re G.M., Jean C.G.;
RT "Androgen-dependent protein from mouse vas deferens. cDNA cloning and
RT protein homology with the aldo-keto reductase superfamily.";
RN J. Biol. Chem. 265:19932-19936(1990).
CC -!- FUNCTION: The role of MVDP in sperm maturation and storage may be
CC related to its potential capacity to produce fructose, or MVDP may
CC play an osmoregulatory role by producing sorbitol.
CC -!- CATALYTIC ACTIVITY: Alditol + NAD(P)(+) = aldose + NAD(P)H.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- INDUCTION: Castration resulted in a marked decrease in the level
CC of the mRNA coding for the protein, whereas administration of
CC testosterone to castrated males resulted in a marked increase.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M81448; AAA39774.1; -.
DR EMBL; J05663; AAA39773.1; -.
DR PIR; A37990; A37990.
DR HSP; P45377; 1FRB.
DR MGD; MGI:101918; Akrlb7.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETRDPASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCTASE_2; 1.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
DR KMOxidoreductase; NADP.
FT INIT MET 0 0 BY SIMILARITY.
FT ACT SITE 48 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
FT CONFLICT 93 93 Q -> D (IN REF. 2).
SQ SEQUENCE 315 AA; 35657 MW; EA3D528294AF8542 CRC64;
Query Match 29.18; Score 503; DB 1; Length 315;
Best Local Similarity 38.78; Pred. No. 2.8e-32;
Matches 123; Conservative 58; Mismatches 121; Indels 16; Gaps 9;
6 TFT-LSNGVKIPGVGFTFSEGSKGSTYTAVTALTGTGRHLDCAWYLNCEGEVGEGR 64

```

QY

Query Match	32.3%;	Score 559.5;	DB 1;	Length 322;
Best Local Similarity	39.3%;	Pred. No. 1.1e-36;		
Matches 127;	Conservative	60;	Mismatches 119;	Indels 17;
				Gaps 6;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 13:59:56 ; Search time 18 Seconds
(without alignments)
940.156 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1731

Sequence: 1 MSNGKFTLSNGVKIGVGF.....MKDTFGYDWPBETAKLSA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	574.5	33.2	312	1	GCY YEAST
2	568	32.8	321	1	YDGF_SCHPO
3	559.5	32.3	322	1	ALDX_SPOSA
4	555	32.1	312	1	YPR1 YEAST
5	547.5	31.6	344	1	ARAL YEAST
6	503	29.1	315	1	ALD1 MOUSE
7	501.5	29.0	310	1	S6PD MALDO
8	499.5	28.9	315	1	ALDR RAT
9	489.5	28.3	284	1	P100 LEIMA
10	488.5	28.2	315	1	ALDR PIG
11	488.5	28.2	315	1	ALDR RABIT
12	487.5	28.2	315	1	ALDR MOUSE
13	487	28.1	316	1	AKBA HUMAN
14	486.5	28.1	315	1	AKBA HUMAN
15	486	28.1	318	1	XYLI PACTA
16	486	28.1	323	1	DHBS MOUSE
17	485	28.0	315	1	ALD2 MOUSE
18	480.5	27.8	301	1	AKB1 MOUSE
19	478	27.6	315	1	ALDR BOVIN
20	478	27.6	324	1	AKAL MOUSE
21	477	27.6	324	1	AKAL HUMAN
22	475.5	27.5	323	1	PE2R RABIT
23	473	27.3	324	1	AKAL RAT
24	467	27.0	324	1	AKAL PIG
25	461	26.6	326	1	AKD1 HUMAN
26	457	26.4	323	1	AKC2 HUMAN
27	456	26.3	295	1	MORA PSBPU
28	454	26.2	323	1	AKC1 HUMAN
29	450.5	26.0	323	1	PGFS BOVIN
30	447.5	25.9	323	1	AKC3 HUMAN
31	445.5	25.7	323	1	PGF2 BOVIN
32	445	25.7	327	1	GRE3 YEAST
33	443.5	25.6	323	1	DBX3 BOVIN

34	439	25.4	329	1	XYLI KLJLA
35	438.5	25.3	323	1	AKC4 HUMAN
36	437	25.2	326	1	AKD1 RAT
37	435.5	25.2	323	1	CRO RANCA
38	432.5	25.0	323	1	CRO RANTE
39	432	25.0	320	1	ALDR HORVU
40	430.5	24.9	323	1	PE2R RAT
41	420.5	24.3	318	1	XYLI PIGST
42	419.5	24.2	322	1	DIDH RAT
43	414.5	23.9	321	1	DTDH MUCMU
44	414	23.9	275	1	DKGA ECOLI
45	413	23.9	282	1	YJ66 YEAST

ALIGNMENTS

RESULT 1
GCY YEAST
ID GCY YEAST STANDARD; PRT; 312 AA.
AC P14065;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GCY protein [EC 1.1.1.-]
GN GCY1 OR GCY OR YOR120W OR O31567 OR YOR3269W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=89005653; PubMed=2901985;
RA Oechsner U., Magdolen V., Bandlow W.;
RT "A nuclear yeast gene (GCY) encodes a polypeptide with high homology to a vertebrate eye lens protein.";
RL FEBS Lett. 238:123-128(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97060020; PubMed=8904341;
RA Wiemann S., Rechmann S., Benes V., Voss H., Schwager C., Vleck C., Stegemann J., Zimmermann J., Effle H., Paces V., Ansoerge W.;
RT "Sequencing and analysis of 51 kb on the right arm of chromosome XV from Saccharomyces cerevisiae reveals 30 open reading frames.";
RL Yeast 12:281-288(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97344368; PubMed=9200815;
RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C., Schwager C., Paces V., Sander C., Ansoerge W.;
RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV.";
RL Yeast 13:655-672(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 25657;
RA Bandlow W.;
RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Unknown function which seems to be not essential.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE FAMILY.
CC VERY SIMILAR TO YEAST YPR1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC
CC EMBL; X13228; CAA31615.1; -.

Search completed: May 28, 2004, 14:06:18
Job time : 21 secs

```
Db      3 TYQLNNGVEIPVLGFGTFKADGE-EAYRAVLEALKAGYRHIDTAAYIQNEESVGOAID 61
QY      66 FLXENSVKREDIFVCTKWNHLHRYEDVLSIDSLKELGLDYVDMFLVHWPPIAAEKNG 125
Db      62 -----SGVPREEMFVTTKLWNSQQTVEQTRQALEKSIKELGLDLYDLIHWPF----- 109
QY      126 QGEPKIGPDGKYVLKDLTENPEP-----TWAMEKIYEDRKARSIGVSNWT 172
Db      110 -----NPKPLRENDANKTRNAEVWRAMEDLYQEGKIRAIQVSNFL 149
QY      173 IADLEKMSFKVMPHANQIEHPFLPNBELVOYCFSKNIMPVAYSPGLSQNVPTTGER 232
Db      150 PHLDALLEATATVPAVNVRLAPGVQDVAYCREKGILLLEAWGPFQO-----GE- 201
QY      233 VSENKTLNEIAEKGNTLAQVLIAGLRGCVVLPKSSNPKRIESNFK--SIELSDADFE 290
Db      202 LFDSKQVQEIQAANHGKSVQAIALAWSLAEGFLPLPKSVTTSRIQANLDCFGIELSHEERE 261
QY      291 AINAVA 296
Db      262 TLKTIA 267

RESULT 15
C98038
conserved hypothetical protein spr1332 [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: C98038
R;Hoskins J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: C98038
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00136.1; PID:g15458977; GSPDB:GN00174
C;Genetics:
A;Gene: spr1332
C;Superfamily: aldehyde reductase

Query Match      29.8%; Score 516; DB 2; Length 280;
Best Local Similarity 35.9%; Pred. No. 7.8e-33;
Matches 110; Conservative 57; Mismatches 83; Indels 56; Gaps 7;

QY      6 TFTLSNGVKIPGVGTFSESGSGGETYATVTTALKTGYRHLDCAWYLNAGEVGEGRD 65
Db      3 TYQLNNGVEIPVLGFGTFKADGE-EAYRAVLEALKAGYRHIDTAAYIQNEESVGOAID 61
QY      66 FLXENSVKREDIFVCTKWNHLHRYEDVLSIDSLKELGLDYVDMFLVHWPPIAAEKNG 125
Db      62 -----SGVPREEMFVTTKLWNSQQTVEQTRQALEKSIKELGLDLYDLIHWPF----- 109
QY      126 QGEPKIGPDGKYVLKDLTENPEP-----TWAMEKIYEDRKARSIGVSNWT 172
Db      110 -----NPKPLRENDANKTRNAEVWRAMEDLYQEGKIRAIQVSNFL 149
QY      173 IADLEKMSFKVMPHANQIEHPFLPNBELVOYCFSKNIMPVAYSPGLSQNVPTTGER 232
Db      150 PHLDALLEATATVPAVNVRLAPGVQDVAYCREKGILLLEAWGPFQO-----GE- 201
QY      233 VSENKTLNEIAEKGNTLAQVLIAGLRGCVVLPKSSNPKRIESNFK--SIELSDADFE 290
Db      202 LFDSKQVQEIQAANHGKSVQAIALAWSLAEGFLPLPKSVTTSRIQANLDCFGIELSHEERE 261
QY      291 AINAVA 296
Db      262 TLKTIA 267
```

```

QY      127 GEPKIGDGYVILXDLTENPEPTWRAMEKIYEDRKARSIGVSNWTTIADLERKSKPAKYM 186
        :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      119 ---SLMPTTEMLTKPDITS---TWKANEALYDSGKARAIGVSFSSKKLTDJLNVARVT 171
        :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      187 PHANQIEIHFFLPNEELVOYCFSKNIMPVASPLGSQNQVPETGE--RVSENKTILNEA 243
        :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      172 PAVNQVECHPVVQQGLHELCKSGVHLSGVSPGQSOK---CEVLKVQLQNPIVEVA 227
        :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      244 EKGNTLAOVLAWGLRGYVVLPKSSNPKRIENFKSIELSDAD--FEAINAVAKGRHF 301
        :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      228 EKLGKTTAQVALRWGLQTGHSVLPKSSSGARLKENLDVFDWSIEDLFTKFSNIPQASSF 287
        :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      302 RFV 304
        :
Db      288 IFL 290

```

```

RESULT 12
T09670
abscisic acid activated protein - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: T09670
R:Deal, M.; Davletova, S.; Oberschall, A.; Messaros, T.; Miskolczi, P.; Torok, K.; Dudits
submitted to the EMBL Data Library, April 1996
A:Description: Calmodulin inhibitors enhance the expression of abscisic acid-activated g
A:Reference number: Z16912
A:Accession: T09670
A:A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-313 <DEA>
A:Cross-references: EMBL:X97606
A:Experimental source: cultivar Regen S; strain RA3
C:Superfamily: aldehyde reductase

```

Query Match	30.2%	Score 523.5;	DB 2;	Length 313;
Best Local Similarity	37.2%;	Pred. No. 2.4e-33;		
Matches 123;	Conservative	61;	Mismatches 106;	Indels 41;
				Gaps 10;

Qy	5	KTTTLNGVKIPGVGFPFASGSGKGYTVATTALKTGYRHLDCAWYYNEGEVGEGR	64
Db	6	KFFQLNTGAKIPSVGLGTWQAS- -FGVAKAVTTAVQVGYRHIDCAEAYKNQSEIGSALK	63
Qy	65	DFLEKNPSPVKREDIFVCTKVMNHLHRYEDVLWMSIDDSLKRLGLDYVDMFLVHWPAAEKN	124
Db	64	K-LCEDGVVKREELWITSKLWCSDDHPEDVPKALDKTLNDLQLOYLDLYLIHWPVS-N-KR	121
Qy	125	GQSEPKIGPDGKVVILKDLTENPE-----PWRAMEKLYEDRKARSIGVSNWITADLEKM	179
Db	122	GTGE-----FWGELNDHADIPSTWKAALGYDSGKAKAYGVSNFSTFKLQDL	168
Qy	180	SKFAKTMPHANQTEIHFFLNFLEELVQYCFPSKNIMEVAYSPLGSONQVPTTGERVSENKTL	239
Db	169	LDVARVPPAVNVOELHPGQQAKLHAFCESKGIHLSGYSPLGSPGVLKSD- -ILKNPVV	225
Qy	240	NETAEKGGNTLAQVLIANGLRGYYVLPKSSNPKREIESNFKSIELSDAD- -FEAI- -	232
Db	226	KETAEKLGKTPGQVLRGGLQAGHSVLLPASTINEARKIKKNLDVYDWSIPEDLFPKFSEIKQ	285
Qy	293	NAVAKGRHFRFVNMKDTFG-----YDVWPEE	318
Db	286	DKLIKGTFF-----VNDYGAFTIBELMDGE	312

RESULT 13
T48:188
aldose reductase-like protein - Arabidopsis thaliana
N:Alternate names: protein F7A7.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
C:Accession: T48188
R:Bavan, M.; Terryn, N.; Argiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De

ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 2000
 A;Reference number: 724487
 A;Accession: T48188
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-320 <BEV>
 A;Cross-references: EMBL;A161946
 A;Experimental source: cultivar Columbia; BAC clone F7A7
 C;Genetics:
 A;Map position: 5
 A;Introns: 48/3; 65/3; 90/2; 119/3; 188/3; 210/3; 291/3
 A;Note: F7A7.190
 C;Superfamily: aldehyde reductase

	Query Match	30.1%;	Score 521;	DB 2;	Length 320;
	Best Local Similarity	35.6%;	Pred. No. 3.8e-33;		
	Matches	117;	Conservative	73;	Mismatches 105; Indels 34; Gaps 10;
Qy	3	NGKFTFLSNGVKIPGVCFGIFASBGSGKGEYITAVTTALKTGYRHLDCAWYVLNBSGEVGE	62		
Db	12	NMESFRLSGHKIPAVGLGTWRS--GSOA-AHAVVTAIVEGGYRHIDTAWEXGDQREVCGQ	69		
Qy	63	IROFLKENPVSQXGEDI EVCVKVMNHLHRYEDVLWSIDDSLKRLGLDYVDMFLVWPIAAE	122		
Db	70	IKRAM--HAGLERRDLEVTSKMCTELSPERVRFALQNTJKEIQLEVLDTLHWPFLRL	127		
Qy	123	KNGGGEPIGPDGKVIKILKOLTE--NPPTWRAMEKIYEDRKARIGVSNWTIADLEKMSK	181		
Db	128	EGASKPPKAG-----DVLDPDMEGVVRMENLKSGLSVRNIGVCFVTVKLNKLLG	178		
Qy	182	FAKVMPHANQIETHPFLPNBELVQYCFSKNIMPVAYSPLGSONQVPTTGERVSENKTLNE	241		
Db	179	FAELIPAVCQMEHPGWRNRDILFECKKNEIH--AYSPLGSQE---GGRDLIHDQTVDR	232		
Qy	242	IAEKGGNLTAQVLTAWGLRRGVVLPKSSNPKRIESNFKSIE--LSADFAEINAVAKGR	299		
Db	233	IACKLNKTPQQLVWKVGLQRTSVIPKSLNPERIKENIKVPDWIIPBQDFQALNSITDQK	292		
Qy	300	HFRFVNKMDTF-----GYDVWPEE	318		
Db	293	--RVIGDEDLFVNKTEGFFSFASVADLWDHE	319		

RESULT 14

oxidoreductase, aldo/keto reductase family SP1478 [imported] - Streptococcus pneumoniae
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: C95172
 R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidelberg, J.D.; Mayhew, M.W.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, E.; Nelson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95172
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-280 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK755572.1; PID:G14972968; GSPDB:GN00164; TIGR:SP48
 A:Experimental source: strain TIGR4

```

Query Match      29.8%; score 516; DB 2; length 280;
Best Local Similarity 35.9%; Pred. No. 7.8e-33;
Matches 110; Conservative 57; Mismatches 83; Indels 56; Gaps 7
Qy      6 TFLISNGKIPGVGFTAGSGKGYTAVTTALKGYRHLDCAWYVLNKEGVGEGRD 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

A;Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
 A;Reference number: Z14677
 A;Accession: T02543
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-290 <ROU>
 A;Cross-references: EMBL:AC004684; NID:G3236234; PID:G3236259
 A;Experimental source: cultivar Columbia
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: H84796
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-290 <STO>
 A;Cross-references: GB:AE002093; NID:G3236259; PIDN:AAC23647.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: F13M22.27; At2g37770
 A;Map position: 2
 A;Introns: 40/3; 57/3; 83/2; 112/3; 204/3
 C;Superfamily: aldehyde reductase

Query Match 30.6%; Score 529; DB 2; Length 290;
 Best Local Similarity 42.3%; Pred. No. 7.9e-34;
 Matches 124; Conservative 51; Mismatches 98; Indels 20; Gaps 8;
 QY 1 MSGKTTFTLSNGVKIPGVGFTGF-ASEGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGE 58
 DB 1 MANAIFFKLNTGAKFVSGLGTWQASPLVGD--AFAAAVGIYRHIDCAIYNEKE 57
 QY 59 VBGIRDFLENPSVKREDIFVCTKVNNHLRYEDVLMSSIDSLKRLGLDYDMFLVHWP 118
 DB 58 IGAVLKK-LFEDRVKREDLFITKSLCTDHDQDQVPEALNRTKDLQLEVDLYLHWP 116
 QY 119 IAAKNGQGFPGDKGYLKLDTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEK 178
 DB 117 ARIKGSVG---IKPENL-----LPVDTPTWKAMEALYDSKARAIGVSNFTKKLAD 167
 QY 179 MSKFAKYMPHANQIEIHPFLPNEELVQYCFKSNIMPVAYSPGSONQVPTTGER--VSEN 236
 DB 168 LLELARVPPAVNQVECHSPWRQTKIQEFCCKGVHLSAYSPGPG--TTWLKSDVLKN 224
 QY 237 KTLNEIAEKGNTLAQVLIANGRLRGYVLPKSNPKRIEENFKSIELSDADF 289
 DB 225 PILNMAEKLKSPQAQVALEWGLQMGHSLVLPKSTNEGRIKENFNVPDWSIPDY 277

RESULT 10
 B84599
 hypothetical protein At2g21260 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C;Accession: B84599
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: B84599
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-309 <STO>
 A;Cross-references: GB:AE002093; NID:G4567261; PIDN:AAD23674.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g21260
 A;Map position: 2
 C;Superfamily: aldehyde reductase
 Query Match 30.58; Score 527.5; DB 2; Length 309;

Best Local Similarity 38.2%; Pred. No. 1.1e-33;
 Matches 120; Conservative 61; Mismatches 118; Indels 15; Gaps 8;
 QY 8 TLSNGVKIPGVGFTGFASEGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGEVGGIRDFL 67
 DB 4 TNSGFKMPIIGLGWRME--KEELRDLIDAIKIGVRLHLDCAANYKNEAEVGEALTEAF 61
 QY 68 KENPSVKREDIFVCTKVNNHLRYEDVLMSSIDSLKRLGLDYDMFLVHWPFAAEKNGQ 127
 DB 62 TTG-LVKREDLFITKLSMSDHGH--VIEACKDSLKQLDYLDFLWHPIATKHTGIG 118
 QY 128 --EPKIGDGYKVKILKLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKSKAKV 185
 DB 119 TDSALGDDG--VIDIDITISLETITWDMKELVMSGLVRSIGISNYDVFTRDCLAYSKI 176
 QY 186 MPHANQIBIHPFLPNEELVQYCFKSNIMPVAYSPGSONQVPTTGERVS--ENKTLNEIA 243
 DB 177 KPAVNIETHFYFQDRLVKFCQKHGICVTAHTPLGGATANAWEFGTVSCLDLDPVKDVA 236
 QY 244 EKGNTLAQVLIANGRLRGYVLPKSNPKRIEENFK--SIELSDADFEAINAVAKGRHP 301
 DB 237 EKYQTVAQIVLRWGIQRTYVIEPTKPERLEENFQVDFQLSKEDMEVKSME--RNY 294
 QY 302 RFVNMKDTFGYDVM 315
 DB 295 RTHQTAKFWGIELY 308
 RESULT 11
 T02542
 probable alcohol dehydrogenase [imported] - Arabidopsis thaliana
 N;Alternate names: aldehyde dehydrogenase homolog F13M22.26
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
 C;Accession: T02542; G84796
 R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.
 submitted to the EMBL Data Library, June 1998
 A;Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
 A;Reference number: Z14677
 A;Accession: T02542
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-294 <ROU>
 A;Cross-references: EMBL:AC004684; NID:G3236234; PID:G3236258
 A;Experimental source: cultivar Columbia
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: G84796
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-294 <STO>
 A;Cross-references: GB:AE002093; NID:G3236259; PIDN:AAC23646.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: F13M22.26; At2g37760
 A;Map position: 2
 A;Introns: 36/3; 53/3; 79/2; 108/3; 200/3
 C;Superfamily: aldehyde reductase
 Query Match 30.2%; Score 523.5; DB 2; Length 294;
 Best Local Similarity 38.3%; Pred. No. 2.2e-33;
 Matches 116; Conservative 61; Mismatches 101; Indels 25; Gaps 7;
 QY 7 FTLSNGVKIPGVGFTGFASEGSGKGTYYTAVTTALKTGYRHLDCAWYLYNEGEVGGIRDF 66
 DB 8 FELNTGAKLPCVGLGTAM-----VATAEQALIKGYRHIDCASIYNEKEIG-GVLKK 60
 QY 67 LKENPSVKREDIFVCTKVNNHLRYEDVLMSSIDSLKRLGLDYDMFLVHWPFAAEKNGQ 126
 DB 61 LIGDGFVKREELFITKLSMSDHLDPVPALEKTLQDLYDLYLHWPASLKE-- 118

aldo-keto reductase (EC 1.-.-.-) YPR1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D9481.8; protein YDR368W
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
C;Accession: S61163; S61563; JCS457; PC4473
R;Ding, H.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of *S. cerevisiae* cosmid 9481.
A;Reference number: S61159
A;Accession: S61163
A;Molecule type: DNA
A;Residues: 1-312 <DIN>
A;Cross-references: EMBL:U28373; NID:g849184; PIDN:AA64804.1; PID:g849189; MIPS:YDR368W
R;Kotani, T.; Juhnke, H.; Sterkel, C.; Zimmermann, F.K.
submitted to the EMBL Data Library, July 1994
A;Description: A new yeast gene with homology to the aldo-keto reductase protein family.
A;Reference number: S61562
A;Accession: S61563
A;Molecule type: DNA
A;Residues: 1-312 <MIO>
A;Cross-references: EMBL:X80642; NID:g1103496; PIDN:CAA56686.1; PID:g1103498
A;Experimental source: strain M5
R;Nakamura, K.; Kondo, S.; Kawai, Y.; Nakajima, N.; Ohno, A.
Biochem. Biotechnol. Biochem. 61, 375-377, 1997
A;Title: Amino acid sequence and characterization of aldo-keto reductase from bakers' yeast
A;Reference number: JCS457; MUID:97212041; PMID:9058981
A;Accession: JCS457
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-312 <NAK>
A;Accession: PC4473
A;Molecule type: protein
A;Residues: 2-36;45-50;68-96, 'X';124-136;178-227;287-293;299-304 <NA2>
C;Genetics:
A;Gene: SGD:YPR1
A;Cross-references: MIPS:YDR368W; SGD:S0002776
A;Map position: 4R
C;Complex: monomer
C;Function:
A;Description: oxidoreductase; NADPH-dependent keto ester reductase; catalyzes the reduction of aldo-keto reductase
C;Superfamily: aldehyde reductase
C;Keywords: monomer; NADP; oxidoreductase
F;2-312/Product: aldo-keto reductase #status experimental <MAT>

Query Match 32.1%; Score 555; DB 2; Length 312;
Best Local Similarity 40.1%; Pred. No. 8.1e-36;
Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;

QY 2 SNGKFTLSNGVKIPGVGFTFASGSGKGTYYTAVTALKTGVRHLDCAWYILNEGEVGE 61
DB 7 NSSATLKLNTGASIPVLGFTGWSVDNNG--YHSVIAALKAGYRHDAAAIYLNEEVGR 64

QY 62 GIRDFLKENSVKREDIFVCTKWNHLYRYEDVLWSDSLKRLGLDYVDMFLVHWPIAA 121
DB 65 AIKD-----SCVPREEIFITKLWGTEQR--DPEALNKSRLGLDYVDMFLVHWVPL 117

QY 122 EKNQGGEPKIGPDGKYVILKDLTE-----NPEPTWRAEKIYEDRKARSIGVSNW 171
DB 118 KTDREV-----TDGNVLICPTLEDGTVIDITKEWNFIKTWELMGELPKTKTGAAGVSNF 171

QY 172 TIADLEK--SKFAKMPHANOIEIHPFLPNEELVQVCFSKNIMPVAYSPLGSONQVPTT 229
DB 172 SINNIKELLESNNKVVPAVQNIIEHLLPQDELIAFCKEIGIVVAYSPPFGSAN-APLL 230

QY 230 GERVENKNTLNEIAKCGNTLAQVLIAGLRRGVVLPKSSNPKRIESNFKSIELSDAPF 289
DB 231 KE-----QAIDNAKKHGVPEAQLIISWSIQRGVYVLAWSNPERIVSNFKIETLPEDDF 285

QY 290 EAINAVAKGRHF-RFVNKK 307
DB 286 KTISSLKSHVGTGTRKRVDMK 304

RESULT 5
S76143
probable aldehyde reductase (EC 1.1.1.-) - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S76143
R;Kaneoko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76143
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-327 <KAN>
A;Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAAL8402.1; PID:gl65348;
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Start codon: GTG
A;Superfamily: aldehyde reductase
C;Keywords: oxidoreductase

Query Match 31.9%; Score 553; DB 1; Length 327;
Best Local Similarity 38.1%; Pred. No. 1.2e-35;
Matches 117; Conservative 69; Mismatches 105; Indels 16; Gaps 6;

QY 5 KFTLSNGVKIPGVGFTFASGSGKGTYYTAVTALKTGVRHLDCAWYILNEGEVSGIR 64
DB 11 KYFLPSNGEQIPALGLGTWKS--SPQVVGQAVEQALDGYRHLDCALYNEAIGATLA 68

QY 65 DFLKENPSVKREDIFVCTKWNHLYRYEDVLWSDSLKRLGLDYVDMFLVHWPIAAEK 124
DB 69 NAFKKG-VKKEELWITSLKWSNAHPDAVLPALKETLQDLGLDYLDLYLHWPEVVI-- 124

QY 125 CGGEPKIG-PDGKVVILKDLTENPEPTWRAEKIYEDRKARSIGVSNWTADLEKSKFA 183
DB 125 ---QPDVGPESGQLLPFTFASLEGTMQALEKAVDLGLCHHIGVSNFSLKLEENLSMA 181

QY 184 KMPHANOIEIHPFLPNEELVQVCFSKNIMPVAYSPLGSONQVPTTGERVSENKTL---- 239
DB 182 RIPFAVQVLEHLYLQSDLLTFANSQNLITATSPLGSGDR-PAAFQQAEPKLLTDPV 240

QY 240 -NETAEKCGNTLAQVLIAGLRRGVVLPKSSNPKRIESNFKSIELSDADPEAINAVAKG 298
DB 241 INGIAEQGCSAAQVLLAWAIQRTVTTPKSNPERLEQNLRADITLTDSMAKIALLD 300

QY 299 RHFREVN 305
DB 301 RHYRYS 307

RESULT 6
T39169
probable oxidoreductase [imported] - fission yeast (*Schizosaccharomyces pombe*)
C;Species: *Schizosaccharomyces pombe*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C;Accession: T50378; T39169
R;Connor, R.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21832
A;Accession: T50378
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-325 <CO2>
A;Cross-references: EMBL:AL021815; PIDN:CAA16997.1; GSPDB:GN00067; SPDB:SPBC8E4.04
A;Experimental source: strain 972h-; cosmid c8E4
C;Genetics:
A;Gene: SPDB:SPAC8E4.04
A;Map position: 2
C;Superfamily: aldehyde reductase

A;Map position: 15R
C;Superfamily: aldehyde reductase
C;Keywords: oxidoreductase

```

Query Match      33.2%; Score 574.5; DB 1; Length 312;
Best Local Similarity 40.8%; Pred. No. 2.4e-37;
Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFILNGVKIPGVGCFEASEGSKGETYATVTTALKGYRHLDCAWVYLNEGVGSGIR 64
Db 10 KILSLNTGAQIPQIGLGTWOSK--ENDAYKAVLTALKDGYRHIDTAAYRNEDOVGQAIK 67
QY 65 DFLKENPSVKREDIFVCTKYWNHHLRYEDVLWSIDSLKRLGLDYDMFLVHWPFAAEKN 124
Db 68 D-----SGVPREEIFVTTKWCITQHSEPV--ALDQSLKRLGLDYDVLMLHWPFARL-- 117
QY 125 GQGEPTGPGGVYILKDLTENPE-----PTFWAMKVIYEDRKARISGV 168
Db 118 -----DPAYIKNEIDIILSVPTKKDGSRAVDITNNFIKTWELMQELPRTGKTKAVGV 168
QY 169 SNMTIADLEKM--SKFAKVMPEHANOIEIHPLEPNEELVQVCFSKNIPVAYSPLGSQNV 226
Db 169 SNFSINLTKOLLASQGNKLTAAANQVSEIHPHLLPODELINFCKSGIYVVEAYSPLGS---- 224
QY 227 PTTGERSYSEKNTLNEIAEKGGNTLTAQVLIAGLARGYVVLFPKSNPKRIENPKFISLD 286
Db 225 --TDAPLLKPEPVILEIAKKNVQSGHVVISWHVQGVVLPKPSVNPDKITNRIFTLST 282
QY 287 ADFEAINAVAK 297
Db 283 EDFEAINNISK 293

```

RESULT 2

Probable oxidoreductase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T38413
 R:Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z21792
 A:Accession: T38413
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-321 <ERO>
 A:Cross-references: EMBL:Z73100; PIDN:CAA97364.1; GSPDB:GN00066; SPDB:SPAC26F1.07
 A:Experimental source: strain 972h-; cosmid c26f1
 C:Genetics:
 A:Gene: SPDB:SPAC26F1.07
 A:Map position: 1
 C:Superfamily: aldehyde reductase

Query Match.	32.8%;	Score 568;	DB 2;	Length 321;
Best Local Similarity	43.0%;	Pred. No. 8.2e-37;		
Matches 131;	Conservative 59;	Mismatches 89;	Indels 26;	Gaps 10;
QY	7	FTLSNGVKYIGVGFTGTFASEGSKGETYATVATKLTGYRHLDCAWYLVNEGEVGEGRD	66	
Db	16	FTIADGSKIPLGLGUTWRSE--PNOTKNAVKTALQGYRHIDAAAIYGNEDVEVDGKHKE	72	
QY	67	LKENPYSVKREDIFVCTKQVNHHLHYEDVLNSIDSLKRLGLDYVDMFLVHWPIAAENXGQ	126	
Db	73	----SGVPRKDIWYTSKLWCNCAHAPAVPKALEKTLKLDLYLDLYLHWVPSF-KTGE	127	
QY	127	GEPIKIGDGKGVILKDLTENP-EPTWAMEKIYEDRKARSIGVGNNTWITADLEKNSKSAKV	185	
Db	128	DKPKDKDGLNLYEK---NPIETWAMEKILETKVYRHLGLSNFNDTLNRLTLVAKV	183	
QY	186	MPHANQTEIHFFLPNEELVQVCFSKNIMPVAYSPLGSON-----QVPTTGBRVSEKNTLN	240	
Db	184	KPAVHQMELHPFLPQTSFEVKHKLGHVTAYSFGNQTNIYESKIP---KLHEHTIQ	239	

```

Qy      241 ETAE-KG-GNTLAQVLIANGLRGGVVLPKSSNPKRITESNFKSIELSDADFEAINAVAKG 298
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      240 KIAKSKGEGVTGATTAWSAIIITRGTSVIPKSVEQIKSNFKYIPLTKEDMDSEINSI--G 297

Qy      299 RHFRF 303
       ||
Db      298 IRARF 302

RESULT 3
S78113
aldehyde reductase (NADPH) (EC 1.1.1.-) - fungus (Sporidiobolus salmonicolor)
N;Alternate names: NADPH-dependent aldehyde reductase
C;Species: Sporidiobolus salmonicolor, Sporobolomyces salmonicolor
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 11-Jun-1999
C;Accession: S78113; S23931
```

RESULT 3

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 14:01:37 ; Search time 20 seconds
(without alignments)
1563.112 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1731

Sequence: 1 MSGNGKFTLLSGVKIPGVGF.....MKDITFDYDWPBETAKNLNSA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	574.5	33.2	312	1 S22846	probable aldehyde
2	568	32.8	321	2 T38413	probable oxidoredu
3	559.5	32.3	323	2 S78113	aldehyde reductase
4	555	32.1	312	2 S61163	aldo-keto reductas
5	553	31.9	327	1 S76143	probable aldehyde
6	548.5	31.7	325	2 T39169	probable oxidoredu
7	547.5	31.6	344	1 S46020	probable aldehyde
8	529.5	30.6	309	2 A84599	hypothetical prote
9	529	30.6	290	2 T02543	aldehyde dehydroge
10	527.5	30.5	309	2 B84599	hypothetical prote
11	523.5	30.2	294	2 T02542	probable alcohol d
12	523.5	30.2	313	2 T09670	absisic acid acti
13	521	30.1	320	2 T48188	aldose reductase-1
14	516	29.8	280	2 C95172	oxidoreductase, al
15	516	29.8	280	2 C98038	conserved hypothet
16	509	29.4	316	2 A37990	aldehyde reductase
17	508.5	29.4	281	2 D86658	oxidoreductase ycg
18	503	29.1	276	2 C70040	plant-metabolite d
19	501.5	29.0	310	2 T17013	D-sorbitol-6-phosp
20	500	28.9	280	2 D69988	plant metabolite d
21	499.5	28.9	316	1 A60603	aldehyde reductase
22	491.5	28.4	350	2 B84797	probable alcohol d
23	490.5	28.3	316	2 A59021	aldehyde reductase
24	490	28.3	274	2 AC1535	oxydoreductases ho
25	488.5	28.2	284	1 A32950	probable aldehyde
26	487.5	28.2	294	2 AC3642	2,5-diketo-D-gluc
27	486.5	28.1	316	1 A39763	aldehyde reductase
28	486	28.1	323	1 A56424	estradiol 17beta-d
29	485	28.0	274	2 D97141	probable aldo/keto

30	485	28.0	316	2 A53440	aldose reductase h
31	484.5	28.0	316	2 I49484	aldehyde reductase
32	483	27.9	274	2 AG1177	oxydoreductases ho
33	480	27.7	278	2 A84131	plant-metabolite d
34	478	27.6	315	1 A35452	aldehyde reductase
35	477	27.6	325	2 A33851	alcohol dehydrogen
36	475.5	27.5	323	1 A45366	20alpha-hydroxyste
37	473	27.3	325	1 JN0629	alcohol dehydrogen
38	466	26.9	316	2 T26766	hypothetical prote
39	465.5	26.9	302	2 A34406	aldehyde reductase
40	461	26.6	326	1 S41120	cholestenone 5beta
41	457	26.4	323	2 JC5240	3alpha-hydroxycho
42	456.5	26.4	295	1 S30383	morphine 6-dehydro
43	456	26.3	323	2 I73676	chlorocone reduct
44	455.5	26.3	329	2 I53872	dihydrodiol dehydr
45	454	26.2	276	2 P83919	plant-metabolite d

ALIGNMENTS

RESULT 1

S22846 probable aldehyde reductase (EC 1.1.1.-) GCY1 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O3269; protein YOR120W; protein YOR3269W

C:Species: Saccharomyces cerevisiae

C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: S22846; S60986; S61678; S67005; S63863

R:Oechsner, U.; Magdalen, V.; Bandlow, W.

FEBS Lett. 238, 123-128, 1988

A:Title: A nuclear yeast gene (GCY) encodes a polypeptide with high homology to a vertebr

A:Reference number: S22846; MUID:89005653; PMID:2901985

A:Accession: S22846

A:Molecule type: DNA

A:Residues: 1-312 <OE>

A:Cross-references: ENBL:X13228; NID:G3737; PIDN:CAA31615.1; PID:G3738

R:Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J.;

submitted to the ENBL Data Library, August 1995

A:Description: Sequencing of 51 kilobases on the right arm of chromosome XV from S. cerev

A:Reference number: S60983

A:Accession: S60986

A:Molecule type: DNA

A:Residues: 1-312 <WIE>

A:Cross-references: ENBL:X90518; NID:G1050808; PIDN:CAA62107.1; PID:G1050812

R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banreivi, A.; Sander, C.; Valencia,

submitted to the ENBL Data Library, December 1995

A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome

A:Reference number: S61643

A:Accession: S61678

A:Molecule type: DNA

A:Residues: 1-312 <HEN>

A:Cross-references: ENBL:X94335; NID:G1262139; PIDN:CAA64040.1; PID:G1164965

R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansorge, W.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66965

A:Accession: S67005

A:Molecule type: DNA

A:Residues: 1-312 <VOS>

A:Cross-references: ENBL:Z75028; NID:G1420316; PIDN:CAA99318.1; PID:G1420317; GSPDB:GN000

A:Experimental source: strain S288C

R:Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J.;

Yeast 12, 281-288, 1996

A:Title: Sequencing and analysis of 51 kb on the right arm of chromosome XV from Sacchar

A:Reference number: S63860; MUID:97060020; PMID:8904341

A:Accession: S63863

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-312 <WIE>

A:Cross-references: ENBL:X90518; NID:G1050808; PIDN:CAA62107.1; PID:G1050812

A:Note: the nucleotide sequence was submitted to the ENBL Data Library, August 1995

C:Genetics: SGD:GCY1; MIPS:YOR120W

A:Gene: SGD:GCY1; MIPS:YOR120W

A:Cross-references: SGD:S0005646; MIPS:YOR120W

Db 85 L-EDGSIKREDLFTTKVW-----PVLWDEVDRLSINSLKALGLEYYVDLLQHWPLCF 136
 QY 122 EK-----NGGEPKIGDPGKYVILKDLTENPEPTWAMEKIY---EDRKA 163
 Db 137 EKIKDPKIGISGLVKTPTVDDSGKTWYAADGYL-----ETYKLEKIYLDPNDRHV 186
 QY 164 RSIGVSNWTIADLEKMSKFAKVMPHANQIEIHPFLPNEELVQVCFKSNIMPVAYSPLGSO 223
 Db 187 RAIGVSNFSIEYLERLIKECRVKPTVQVETHPHLPQWELRKFCFMDHILLTAYSPLGSH 246
 QY 224 N-----QVPTTGERVSEKNTLEIAEKGGNTLAQVLIANGLRGGYVVLPKSNPKRIESNF 279
 Db 247 GAPNLKIP-----LVYKLAEKYNTVGNLLISYHROGTIVIPRSLNPRVRISSI 296
 QY 280 KSIELSDADFEAINAVAKGSHFRFVN 305
 Db 297 EFASLTQDELQELNDFGKYPVRID 322

RESULT 14
 AAU76345
 ID AAU76345 standard; protein; 344 AA.
 XX AC AAU76345;
 XX DT 21-MAY-2002 (first entry)
 XX DE Yeast D-arabinose dehydrogenase (ARA).
 XX KW Ascorbic acid; vitamin C; scurvy; recombinant yeast; enzyme;
 XX KW L-galactose dehydrogenase; LGDH; L-galactono-1,4-lactone dehydrogenase;
 XX KW AGD; D-arabinose dehydrogenase; ARA; D-arabinono-1,4-lactone oxidase;
 XX KW ALD; L-gulono-1,4-lactone oxidase; GLO; aldono-lactonase; AL.
 XX OS Saccharomyces cerevisiae.
 XX PN WO200210425-A2.
 XX PD 07-FEB-2002.
 XX PF 02-AUG-2001; 2001WO-GB003485.
 XX PR 02-AUG-2000; 2000US-00630983.
 XX PA (BIOP-) BIPOLO SCARL.
 XX FA (WHALE/) WHALLEY K.
 XX PI Porro D, Sauer M;
 XX PN WPI; 2002-217125/27.
 XX DR N-PSDB; ABK10132.
 XX PT Generating ascorbic acid or its salt, involves culturing yeast capable of
 PT converting ascorbic acid precursor into ascorbic acid in medium
 PT comprising ascorbic acid precursor, and isolating ascorbic acid.
 XX PS Claim 12; Page 90-91; 95pp; English.
 XX CC The invention relates to generating ascorbic acid or its salt, involves
 CC obtaining a recombinant yeast capable of converting an ascorbic acid
 CC precursor into ascorbic acid, culturing the recombinant yeast in a medium
 CC comprising an ascorbic acid precursor, thus forming ascorbic acid, and
 CC isolating the ascorbic acid. Also include are stabilising ascorbic acid
 CC or its salt in a medium, by culturing a yeast in a medium comprising
 CC ascorbic acid or its salt and a recombinant yeast functionally
 CC transformed with a coding region encoding a protein having an enzyme
 CC activity selected from L-galactose dehydrogenase (LDH), L-galactono-1, 4
 CC -lactone dehydrogenase (AGD), D-arabinose dehydrogenase (ARA), D-
 CC -lactone dehydrogenase (AGD), L-galactono-1,4-lactone oxidase (GLO)
 CC and aldono-lactonase (AL) activity, where the recombinant yeast is capable
 CC of converting to ascorbic acid at least about 25% ascorbic acid precursor
 CC or is capable of producing at least 20 mg ascorbic acid/L medium, when

CC the yeast is cultured in a medium comprising one ascorbic acid precursor.
 CC The ascorbic acid produced (Vitamin C) is a powerful antioxidant, a
 CC deficiency of which causes scurvy in humans. The present sequence
 CC represents yeast ARA
 XX Sequence 344 AA;
 SQ Query Match 31.6%; Score 547.5; DB 5; Length 344;
 Best Local Similarity 37.4%; Pred. No. 9.7e-45;
 Matches 122; Conservative 58; Mismatches 91; Indels 55; Gaps 8;
 QY 7 FTLSNGVKIPGVGTFPASGSGKGTETVATVTTALKTYGRHLDCAWYVINEGEVGEGRIDF 66
 Db 25 FSLNNGVRIPALGNTANPHEKLAETKQAVKAAIKAGYRHIDTAWAYETEFVGEAIKEL 84
 QY 67 LKENPSVKREDIFVCTKVNNHLHRYEDVLM-----SIDDSLKRLGLDYVDMFLVHWFLAA 121
 Db 85 L-EDGSIKREDLFTTKVW-----PVLWDEVDRLSINSLKALGLEYYVDLLQHWPLCF 136
 QY 122 EK-----NGGEPKIGDPGKYVILKDLTENPEPTWAMEKIY---EDRKA 163
 Db 137 EKIKDPKIGISGLVKTPTVDDSGKTWYAADGYL-----ETYKLEKIYLDPNDRHV 186
 QY 164 RSIGVSNWTIADLEKMSKFAKVMPHANQIEIHPFLPNEELVQVCFKSNIMPVAYSPLGSO 223
 Db 187 RAIGVSNFSIEYLERLIKECRVKPTVQVETHPHLPQWELRKFCFMDHILLTAYSPLGSH 246
 QY 224 N-----QVPTTGERVSEKNTLEIAEKGGNTLAQVLIANGLRGGYVVLPKSNPKRIESNF 279
 Db 247 GAPNLKIP-----LVYKLAEKYNTVGNLLISYHROGTIVIPRSLNPRVRISSI 296
 QY 280 KSIELSDADFEAINAVAKGSHFRFVN 305
 Db 297 EFASLTQDELQELNDFGKYPVRID 322

RESULT 15
 ABP53552
 ID ABP53552 standard; protein; 328 AA.
 XX AC ABP53552;
 XX DT 16-DEC-2002 (first entry)
 XX DE T. megachiliensis erythrose reductase type I SEQ ID NO:3.
 XX KW Erythrose reductase; enzyme; erythritol; D-erythrose.
 XX OS Trichosporonoides megachiliensis.
 XX PN EP1221478-A2.
 XX PD 10-JUL-2002.
 XX PF 04-JAN-2002; 2002EP-00000321.
 XX PR 09-JAN-2001; 2001JP-00001294.
 XX PR 18-DEC-2001; 2001JP-00384357.
 XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.
 XX PN (NIKM) NIKKEN CHEM CO LTD.
 XX PI Ookura T, Kasumi T, Asaba E;
 XX WPI; 2002-659435/71.
 XX DR N-PSDB; ABQ82227.
 XX PT Novel protein having erythrose reductase activity such as erythrose
 PT reductase type I, II or III, useful for production of erythritol on an
 PT industrial scale.
 XX PS Claim 17; Page 24-26; 39pp; English.
 XX

XX PN WO200264766-A2.
XX PD 22-AUG-2002.
XX PF 21-DEC-2001; 2001WO-EP015398.
XX PR 22-DEC-2000; 2000EP-00870318.
XX PR 04-JAN-2001; 2001EP-00870002.
XX PR 09-JAN-2001; 2001EP-00870003.
XX PA (JANC) JANSSEN PHARM NV.
XX PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX DR WPI; 2002-667002/71.
XX DR N-PSDB; ABQ76569.
XX PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
XX PT medicament for treating, preventing and/or alleviating yeast or fungal
XX PT infections or proliferative disorders, or for preventing apoptosis in
XX PT certain diseases.
XX PS Claim 36; Fig 2; 344pp; English.
XX CC This invention describes a novel nucleic acid representing a synthetic
XX CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
XX CC resistant yeast or fungi, identifying, or obtaining and identifying
XX CC Candida spp. sequences that are differentially expressed in a pathway
XX CC eventually leading to programmed cell death or identifying inhibitors or
XX CC inhibitor sequences of Bax-induced cell death. The products of the
XX CC invention have cytostatic, fungicide; immunosuppressive, virucide and
XX CC vasotropic activity and can be used in vaccines or for gene therapy. The
XX CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
XX CC antisense molecules and antibodies are useful as medicaments or in
XX CC preparing a medicament for treating, preventing and/or alleviating
XX CC diseases associated with yeast or fungi or proliferative disorders, such
XX CC as cancer, or for preventing apoptosis in certain diseases. The compounds
XX CC or polypeptides, or the genetically modified organism are useful for
XX CC preparing a medicament for modifying the endogenous flora of humans and
XX CC other mammals. The vaccine is useful for immunising against yeast or
XX CC fungal infections. Apoptosis-related diseases include autoimmune disease,
XX CC ischaemia, diseases related with viral infections or neurodegenerations.
XX CC This sequence represents a polypeptide associated with the Bax gene
XX CC described in the disclosure of the invention
XX SQ Sequence 295 AA;
Query Match 31.8%; Score 550; DB 5; Length 295;
Best Local Similarity 41.9%; Pred. No. 4.4e-45;
Matches 125; Conservative 56; Mismatches 87; Indels 30; Gaps 8;
QY 3 NGKFTLSNGVKIPGVGFTFASGSGKGTAVTTALTGTGYRHLDCAWYLNVEGVGEG 62
Db 8 NDTFTLNNKNI FAVGLGTWQAT-NEDEAYRAVLALKNKYHIDTAALYGNVEGVGKA 66
QY 63 IRDFLKNPSYKREDIFVCTKVMNHLHRYEDVLWSDSDSLKRLGLDYVDMFLVHPVIAAE 122
Db 67 IKD-----SGVPRBELFVTKLWNADHK--NIEBALETSKLKGLNVDLYLHWPSAD 119
QY 123 KNGGGEKIPGDKYVILKOLTENPBTWRAMEKIYED-RKASIGVSNWTIADLEKM-- 179
Db 120 KSTN---KPYTDFDYV-----DTRYGLQVYKNSKKIRAIQVSNFTKKLERLLS 166
QY 180 SKFAKVMFANQIETHPFLNEELVQYCFGKNIMFVAYSPIGSONQVPTTGERVSENKTL 239
Db 167 SEGVDVVPVANNQIEAHPLLTQPELYDLKKEGIVLEAYSPLGS-----INSLFKNETI 220
QY 240 NEIAEKGNLTQAVLLIANGLRGCVVLPKSNPNFIESNFKSIELSDADPEANAVAK 297
Db 221 VKIAEKNGVEPAQVLVSWAIQKRWLPKSVTESRVSINLKTFTLPSEDFETLNKUSE 278

RESULT 13
ABG93050
ID ABG93050 standard; protein; 344 AA.
XX AC ABG93050;
XX DT 21-NOV-2002 (first entry)
XX DE S. cerevisiae BAX-associated protein fragment SEQ ID 58.
XX KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
XX KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
XX KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
XX KW neurodegeneration; cell death.
XX OS Saccharomyces cerevisiae.
XX PN WO200264766-A2.
XX PD 22-AUG-2002.
XX PF 21-DEC-2001; 2001WO-EP015398.
XX PR 22-DEC-2000; 2000EP-00870318.
XX PR 04-JAN-2001; 2001EP-00870002.
XX PR 09-JAN-2001; 2001EP-00870003.
XX PA (JANC) JANSSEN PHARM NV.
XX PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX DR WPI; 2002-667002/71.
XX DR N-PSDB; ABQ76316.
XX PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
XX PT medicament for treating, preventing and/or alleviating yeast or fungal
XX PT infections or proliferative disorders, or for preventing apoptosis in
XX PT certain diseases.
XX PS Claim 36; Fig 1; 344pp; English.
XX CC This invention describes a novel nucleic acid representing a synthetic
XX CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
XX CC resistant yeast or fungi, identifying, or obtaining and identifying
XX CC Candida spp. sequences that are differentially expressed in a pathway
XX CC eventually leading to programmed cell death or identifying inhibitors or
XX CC inhibitor sequences of Bax-induced cell death. The products of the
XX CC invention have cytostatic, fungicide; immunosuppressive, virucide and
XX CC vasotropic activity and can be used in vaccines or for gene therapy. The
XX CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
XX CC antisense molecules and antibodies are useful as medicaments or in
XX CC preparing a medicament for treating, preventing and/or alleviating
XX CC diseases associated with yeast or fungi or proliferative disorders, such
XX CC as cancer, or for preventing apoptosis in certain diseases. The compounds
XX CC or polypeptides, or the genetically modified organism are useful for
XX CC preparing a medicament for modifying the endogenous flora of humans and
XX CC other mammals. The vaccine is useful for immunising against yeast or
XX CC fungal infections. Apoptosis-related diseases include autoimmune disease,
XX CC ischaemia, diseases related with viral infections or neurodegenerations.
XX CC This sequence represents a polypeptide associated with the Bax gene
XX CC described in the disclosure of the invention
XX SQ Sequence 344 AA;
Query Match 31.6%; Score 547.5; DB 5; Length 344;
Best Local Similarity 37.4%; Pred. No. 9.7e-45;
Matches 122; Conservative 58; Mismatches 91; Indels 55; Gaps 8;
QY 7 FTLSNGVKIPGVGFTFASGSGKGTAVTTALTGTGYRHLDCAWYLNVEGVGEGIDF 66
Db 25 FSLNNGVRIPALGLGTANTPEKLAETKQAVKAAIKAGYRHDITAWAYETFPVGEAIKEL 84
QY 67 LKENPSVKREDIFVCTKVMNHLHRYEDVLW-----SIDDSLKRLGLDYVDMFLVHPVIAA 121

PT Carbonyl reductase gene and protein - used in the production of R-gamma-
 PT substituted-beta-hydroxybutyrate, for use in drugs and agrochemicals.
 XX PS Claim 1; Page 9-10; 14pp; Japanese.
 XX The present sequence is that of a carbonyl reductase (ALD) isolated from
 CC Sporobolomyces salmonicolor IF01038. The enzyme is used in the synthesis
 CC of an (R)-gamma-substituted-beta-hydroxybutyrate (GSBH) from a gamma-
 CC substituted acetoacetate. The GSBH is useful as a raw material for the
 CC synthesis of drugs and agricultural chemicals. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX SQ Sequence 323 AA;

Query Match 32.3%; Score 559.5; DB 2; Length 323;
 Best Local Similarity 39.3%; Pred. No. 5.8e-46;
 Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
 QY 6 TFLTSLNGVKIPGVGFGTFASEGSGKGETYTAVTTALKTGYRHLDCAWYLNAGEVGEGRD 65
 DB 4 TTTLNTGASLELVGYGTW--QAAPGEVGGQGVKVAIETGYRHLDLAKVYSNQPEVGAAIKE 61
 QY 66 FLKPNPSVKREDIFVCTKWVNHHRVEDVLSDDSLKRLGLDYVDMFLVHWPPIAEKNG 125
 DB 62 -----AGVKKREDLFTSKLWNSHRPEQVEPALDDTLKELGLDYLDLYLHWPVAPFPEG 116
 QY 126 QGEPKIGPDGKIVLK-DLTENPEPTWRAMEKIYEDRKARSIGVSNWTTIADLEKMSKFAK 184
 DB 117 DITQNLFPKANDKEVKLDLEVSLVDTWKAMVLLDTGKVAIGVSNFDAQWDAIIEATG 176
 QY 185 VMPHANOIETHPLPNEELVOYCFSKNIMPVAYSPILGSONQVPTTG-ERVSENKTLNEIA 243
 DB 177 VTFSVNOIERHPLLOPELLIAHAKKNIHITAYSPLGN-----TVGAPLLVQHPFKRIA 232
 QY 244 EKGNTLAQVLIAGLRGVVLPKSNPKRIESNFKSIELSADFEAINAVAKGRHFRF 303
 DB 233 EKGCTPAQVLIAGVGGSHVIPKSVTPSRIGENFKQVLSQEDVDVSKLGEVGGRRR 292
 QY 304 VNMKDTFG----YDWVPEETAKN 322
 DB 293 YNIPCTYSPKWDINVFGEDEKS 315

RESULT 11
 AAG63564
 ID AAG63564 standard; protein; 313 AA.

AC AAG63564;
 XX 15-OCT-2001 (first entry)
 DE Amino acid sequence of a ketone reductase.
 XX Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding; ketone reductase.
 XX Saccharomyces cerevisiae.

XX WO20015342-A2.
 XX 02-AUG-2001.
 XX 31-JAN-2001; 2001WO-US003186.
 XX 31-JAN-2000; 2000US-00494921.
 XX 08-DEC-2000; 2000US-00734237.
 XX (BIOC-) BIOCATALYTICS INC.
 XX Rozzell DJ, Bui P, Hua L;
 XX WPI; 2001-483235/52.
 XX N-PSDB; AAH74581.

XX Designing synthetic nucleic acid sequences for improved amplification,
 PT expression in host cell, by comparing free energy of folding of a
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement.
 XX Claim 6; Page 69-70; 117pp; English.

XX The present sequence represents a ketone reductase. The polynucleotide
 CC sequence was modified using the method of the invention. The
 CC specification describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with a
 CC different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotide, and comparing
 CC this with that of the original polynucleotide. The method is useful for
 CC developing nucleic acid sequences that enhance expression of the encoded
 CC protein in a heterologous host. The design and preparation of the
 CC synthetic genes are used in application of gene shuffling, directed
 CC evolution and molecular breeding methods. The method allows expression of
 CC genes from various organisms such as mammals, plants, fungi and
 CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
 CC hosts at commercially viable levels, in particular proteins with low
 CC yield such as methionine gamma-lyase from *P. putida*

XX SQ Sequence 313 AA;

Query Match 32.1%; Score 555; DB 4; Length 313;
 Best Local Similarity 40.1%; Pred. No. 1.5e-45;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 QY 2 SNGKTFLLSNGVKIPGVGFGTFASEGSGKGETYTAVTTALKTGYRHLDCAWYLNAGEVGE 61
 DB 8 NSSATLKLNTGASIPVLGFGTWRSDNNG--YHSVIAALKAGYRHIDAAAIYLNEEBVG 65
 QY 62 GIRDPLKENPSVKREDIFVCTKWVNHHRVEDVLSDDSLKRLGLDYVDMFLVHWPPIAA 121
 DB 66 AIKD-----SGVPREEIPITTKLWGTFQR--DPAALNKLKRLGLDYVDLYLHWPVPL 118
 QY 122 EKNQGEPEKIGPDGKIVLKDLTE-----NPEPTWRAMEKIYEDRKARSIGVSNW 171
 DB 119 KTDV-----TDGNVLCTPLEDGTVDIDTKENWFIKTWELMQELPKTKTKAVGVSNF 172
 QY 172 TIADLEKM--SKFAKMPHANOIETHPLPNEELVOYCFSKNIMPVAYSPILGSONQVPTT 229
 DB 173 SINNIKELLESFNKKVWPATNQIEIHPLLOPELLIAFCKEKGIVVEAYSPFGSAN-APLL 231
 QY 230 GERVSENKTLNEIAEKGGNTLAQVLIAGLRGVVLPKSNPKRIESNFKSIELSADFE 289
 DB 232 KE-----QAIDMAKKGVEPAQLIISWSIQRGVVLAKSVNPERIVSNFKIFTLPEDDF 286
 QY 290 BAINAVAKGRHF-RPVNMK 307
 DB 287 KTISNLSKVHGTGRVDMK 305

RESULT 12
 AAG93303
 ID AAG93303 standard; protein; 295 AA.
 XX AAG93303;
 XX 21-NOV-2002 (first entry)
 DE C. albicans BAX-associated protein fragment SEQ ID 564.
 XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death.
 XX Candida albicans.

QY 227 PTTGERSVSEKTLNEIAEKGNLTIAQVLIANGLRGCVVLPKSSNPKRIESNFKSIELSD 286
 Db 225 --TDAPLLKEPVLIEIAKNNVQGHVVISWHVQGVVLPKSNVDPRIKTRKIFTLST 282
 QY 287 ADFAINAVAK 297
 Db 283 EDFAINNISK 293

RESULT 7
 AAG63566
 ID AAG63566 standard; protein; 313 AA.
 AC AAG63566;
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE Synthetic amino acid sequence of a yeast protein.
 XX
 KW Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding.
 XX
 OS Synthetic.
 OS Saccharomyces cerevisiae.
 XX
 PN W0200155342-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US003186.
 XX
 PR 31-JAN-2000; 2000US-00494921.
 PR 08-DEC-2000; 2000US-00734237.
 XX
 PA (BIOC-) BIOCATALYTICS INC.
 XX
 PI Rozzell DJ, Bui P, Hua L;
 XX
 DR WPI; 2001-483235/52.
 DR N-PSDB; AAH74584.
 XX
 PT Designing synthetic nucleic acid sequences for improved amplification,
 PT expression in host cell, by comparing free energy of folding of a
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement.
 XX
 PS Claim 6; Page 73-74; 117pp; English.
 XX
 CC The present sequence is a modified yeast protein. The polynucleotide
 CC sequence was modified using the method of the invention. The
 CC specification describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with a
 CC different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotide, and comparing
 CC this with that of the original polynucleotide. The method is useful for
 CC developing nucleic acid sequences that enhance expression of the encoded
 CC protein in a heterologous host. The design and preparation of the
 CC synthetic genes are used in application of gene shuffling, directed
 CC evolution and molecular breeding methods. The method allows expression of
 CC genes from various organisms such as mammals, plants, yeast, fungi and
 CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
 CC hosts at commercially viable levels, in particular proteins with low
 CC yield such as methionine gamma-lyase from P. putida
 XX
 SQ Sequence 313 AA;
 Query Match 33.28; Score 574.5; DB 4; Length 313;
 Best Local Similarity 40.88; Pred. No. 1.9e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KFTLSNGVKIPGVGTFPASEGSKGETYTAVTTALTGYRHLDCAWYILNEGEVGEIR 64
 Db 11 KILSLNTGAQIPQIGLTWQSK--ENDAYKAVLTALKDGYRHIDTAALYRNEDQVGOAIK 68
 QY 65 DFLKENSVRKREDIFVCTKWNHLHRYEDVLSDDSLKGLDYDMFLVHWPIAAEKN 124
 Db 69 D-----SGVPREEIFVTTKLWCTQHPEV--ALDQSLKRLGLDYDLYLHWHPARL--- 118
 QY 125 GQGEPKIGDCKYVILKDLTENPE-----PTWEAMEKIYEDRKARSIGV 168
 Db 119 -----DPAYINEDILSVPTKDGSRADITNNFIKTWELMOELPKTKTKAVGV 169
 QY 169 SNWTIADLEKM--SKFAKVMPHANQIEIHPLPNEELVQYCFSKNIMPVAYSPIGSQNV 226
 Db 170 SNFSINNLLKDLASQGNKLTAAANQVBIHPLLPQDELINFCKSGKIVVEAYSPLGS--- 225
 QY 227 PTTGERSVSEKTLNEIAEKGNLTIAQVLIANGLRGCVVLPKSSNPKRIESNFKSIELSD 286
 Db 226 --TDAPLLKEPVLIEIAKNNVQGHVVISWHVQGVVLPKSNVDPRIKTRKIFTLST 283
 QY 287 ADFAINAVAK 297
 Db 284 EDFAINNISK 294

RESULT 8
 AAW29218
 ID AAW29218 standard; protein; 312 AA.
 XX
 AC AAW29218;
 XX
 DT 02-MAR-1998 (first entry)
 XX
 DE S. cerevisiae uronate dehydrogenase variant R57E.
 XX
 KW uronate reductase; uronate dehydrogenase; Saccharomyces cerevisiae;
 KW synthesis; L-ascorbic acid; variant.
 XX
 OS Saccharomyces cerevisiae.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 57 /label= R57E
 FT /note= "wild type arginine replaced by glutamine"
 XX
 PN DE19604798-A1.
 XX
 PD 14-AUG-1997.
 XX
 PF 09-FEB-1996; 96DE-01004798.
 XX
 PR 09-FEB-1996; 96DE-01004798.
 XX
 PA (HERB-) HERBSTREITH & FOX PEKTIN-FAB NEUENBUERG.
 XX
 PI Mattes R, Kulbe K;
 XX
 DR WPI; 1997-403701/38.
 XX
 PT Recombinant uronate reductase and L-gulonono-gamma-lactone oxidase - used
 PT in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 XX
 PS Claim 31; Page; 26pp; German.
 XX
 CC The present sequence represents a uronate dehydrogenase (reductase) R57E
 CC variant. The variant can be used for synthesis of L-ascorbic acid. DNA
 CC encoding L-gulonono-gamma-lactone oxidase (GluOx) is also used in the
 CC production of L-ascorbic acid. The uronate dehydrogenase has higher
 CC affinity for substrate (D-glucuronic or galacturonic acids) and coenzymes
 CC than the similar enzyme from Lipomyces starkeyi, also better stability
 CC and greater yield when produced recombinantly. NB. This sequence was
 CC created using the wild type uronate dehydrogenase sequence given in
 CC Figure 3 of the specification

PA (BIOC-) BIOCATALYTICS INC.

PI Rozzell DU, Bui P, Hua L;

DR WPI; 2001-493235/52.

DR N-PSDB; AAH74583.

XX Designing synthetic nucleic acid sequences for improved amplification,
PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.

PS Claim 6; Page 72-73; 117pp; English.

XX The present sequence represents a yeast protein. The polynucleotide
CC sequence was modified using the method of the invention. The
CC specification describes a method for designing a synthetic
CC polynucleotide. The method comprises providing a starting polynucleotide,
CC determining the predicted free energy of folding per base of the
CC polynucleotide, modifying the polynucleotide by replacing a codon with a
CC different codon to provide a modified polynucleotide, determining free
CC energy of folding per base of the modified polynucleotide, and comparing
CC this with that of the original polynucleotide. The method is useful for
CC developing nucleic acid sequences that enhance expression of the encoded
CC protein in a heterologous host. The design and preparation of the
CC synthetic genes are used in application of gene shuffling, directed
CC evolution and molecular breeding methods. The method allows expression of
CC genes from various organisms such as mammals, plants, yeast, fungi and
CC bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic
CC hosts at commercially viable levels, in particular proteins with low
CC yield such as methionine gamma-lyase from *P. putida*

XX Sequence 312 AA;

Query Match 33.2%; Score 574.5; DB 4; Length 312;

Best Local Similarity 40.8%; Pred. No. 1.9e-47;

Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFILNSGVKIPGVGFTFASGSGKGETYATVTTALKTGYRHLDCAWYILNEGEVGEIGR 64

DB 10 KILSLNTGAQIPQIGLGTWQSK--ENDAYKAVLTALKDGYRHDTAAYRNEDQVQGAIK 67

QY 65 DFLKENSVKREDIFVCTKVNHLHRYEDVLSWSDSLKRLGLDYDMFLVHPIAAEKV 124

DB 68 D-----SGVPREEIFVTIKLWCTQHHEPEV--ALDQSLKRLGLDYDMFLVHPIAAEKV 117

QY 125 GQGEPIKPGDGKGVILKDLTENPE-----PTWRAMEKIVYEDRKARSIGV 168

DB 118 -----DPAYIKNEDILSVPTKDGSRVADITNNWFIKTWELMQELPKTKTKAVGV 168

QY 169 SNWTIADLEKM--SKFAKMPHANQIETHPFLPNEELVQYCFKSNIMPVAYSPGQSNQV 226

DB 169 SNFSINNKLKLLASQGNKLTAAQVQVEIHPLLPQDELINFCCKSGIIVVEAYSPLGS---- 224

QY 227 PTTGERSVENKTLNEIAEGKGNLTAAQVLIAGLRGYYVVLKPSNPKARIESNFKSIELSD 296

DB 225 --TDAPIKPEVILEIAKNNVQCHVVISWHVQRGVIVLPLKSVNPDRIKINRKITLTST 292

QY 287 ADFAINAVAK 297

DB 283 EDFAINNISK 293

RESULT 6

ID ABG93198

XX ABG93198 standard; protein; 312 AA.

AC ABG93198;

XX 21-NOV-2002 (first entry)

DT S. cerevisiae BAX-associated protein fragment SEQ ID 354.

DE

XX

KW

KW

KW

KW

OS

XX

PN

XX

XX

PD

XX

XX

PF

XX

XX

PR

XX

PR

XX

PR

XX

PA

XX

PI

XX

DR

XX

DR

XX

XX

PT

XX

PT

XX

PT

XX

XX

PS

XX

CC

CC

CC

CC

CC

CC

Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
neurodegeneration; cell death.

Saccharomycetes cerevisiae.

WO200264766-A2.

22-AUG-2002.

21-DEC-2001; 2001WO-EP015398.

22-DEC-2000; 2000EP-00870318.

04-JAN-2001; 2001EP-00870002.

09-JAN-2001; 2001EP-00870003.

(JANC) JANSSEN PHARM NV.

Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;

WPI; 2002-667002/71.

N-PSDB; ABQ76464.

New isolated nucleic acid representing a synthetic BAX-gene, useful as
medicament for treating, preventing and/or alleviating yeast or fungal
infections or proliferative disorders, or for preventing apoptosis in
certain diseases.

Claim 36; Fig 1; 344pp; English.

This invention describes a novel nucleic acid representing a synthetic
Bax gene. The Bax gene of the invention is useful for identifying Bax-
resistant yeast or fungi, identifying, or obtaining and identifying
Candida spp. sequences that are differentially expressed in a pathway
eventually leading to programmed cell death or identifying inhibitors or
inhibitor sequences of Bax-induced cell death. The products of the
invention have cytostatic, fungicide, immunosuppressive, virucide and
vasotropic activity and can be used in vaccines or for gene therapy. The
isolated nucleic acids, polypeptides, pharmaceutical compositions,
antisense molecules and antibodies are useful as medicaments or in
preparing a medicament for treating, preventing and/or alleviating
diseases associated with yeast or fungi or proliferative disorders, such
as cancer, or for preventing apoptosis in certain diseases. The compounds
or polypeptides, or the genetically modified organism are useful for
preparing a medicament for modifying the endogenous flora of humans and
other mammals. The vaccine is useful for immunising against yeast or
fungal infections. Apoptosis-related diseases include autoimmune disease,
ischaemia, diseases related with viral infections or neurodegenerations.
This sequence represents a polypeptide associated with the Bax gene
described in the disclosure of the invention

Sequence 312 AA;

Query Match 33.2%; Score 574.5; DB 5; Length 312;

Best Local Similarity 40.8%; Pred. No. 1.9e-47;

Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFILNSGVKIPGVGFTFASGSGKGETYATVTTALKTGYRHLDCAWYILNEGEVGEIGR 64

DB 10 KILSLNTGAQIPQIGLGTWQSK--ENDAYKAVLTALKDGYRHDTAAYRNEDQVQGAIK 67

QY 65 DFLKENSVKREDIFVCTKVNHLHRYEDVLSWSDSLKRLGLDYDMFLVHPIAAEKV 124

DB 68 D-----SGVPREEIFVTIKLWCTQHHEPEV--ALDQSLKRLGLDYDMFLVHPIAAEKV 117

QY 125 GQGEPIKPGDGKGVILKDLTENPE-----PTWRAMEKIVYEDRKARSIGV 168

DB 118 -----DPAYIKNEDILSVPTKDGSRVADITNNWFIKTWELMQELPKTKTKAVGV 168

QY 169 SNWTIADLEKM--SKFAKMPHANQIETHPFLPNEELVQYCFKSNIMPVAYSPGQSNQV 226

DB 169 SNFSINNKLKLLASQGNKLTAAQVQVEIHPLLPQDELINFCCKSGIIVVEAYSPLGS---- 224

DR WPI; 1997-403701/38.
 XX
 PT Recombinant uronate reductase and L-gulono-gamma-lactone oxidase - used
 PT in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 XX
 XX Claim 31; Page; 26pp; German.
 XX
 CC The present sequence represents a uronate dehydrogenase (reductase) Q29G
 CC variant. The variant can be used for synthesis of L-ascorbic acid. DNA
 CC encoding L-gulono-gamma-lactone oxidase (GluOx) is also used in the
 CC production of L-ascorbic acid. The uronate dehydrogenase has higher
 CC affinity for substrate (D-glucuronic or galacturonic acids) and coenzymes
 CC than the similar enzyme from *Lipomyces starkeyi*, also better stability
 CC and greater yield when produced recombinantly. NB. This sequence was
 CC created using the wild type uronate dehydrogenase sequence given in
 CC Figure 3 of the specification
 XX
 XX Sequence 312 AA;
 SQ

Query Match 33.2%; Score 575.5; DB 2; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.5e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFTLSNGVKIPGVGFTFASGSKGETYTAVTALKTGYRHLDCAWYVINEGEVGEIR 64
 Db 10 KILSLNTGAQIPQIGLGTWQSK--ENDAYKAVLTALKDGYRHIDTAIYRNEDQVGOAIK 67
 QY 65 DFLKENSVPKREDIFVCTKWNHLHRYEDVLMISDLSKRLGLDYDMFLVHWPAAEKN 124
 Db 68 D-----SGVPREEIFVTIKLWCTQHHEPEV--ALDQSKRLGLDYDVLVHWPAAEL--- 117
 QY 125 GQGEPKIGPDKGVILKDLTENPE-----PTWRAMEKIVEDRKARSIGV 168
 Db 118 -----DPAYIKNEDILSVPTKDGSRADITNWNFIKTWELMQELPKTKTKAVGV 168
 QY 169 SNWTIADLEKM--SKFAKVMPHANQIEIHPPLNEELVOYCFCKNIMPVAYSPLGSONOV 226
 Db 169 SNFSINNKLKOLLASQGNKLTAAANQVEIHPLLPQDELINFCCKSGIWEAYSPLGS---- 224
 QY 227 PTTGERSVENKTLNIEAEKGNLTAAQVLIAGLRGYYVLPKSNPKRIESNFKSIELSD 286
 Db 225 --TDAPLKEPVEILEIAKNNVQPGHVVISHWVQGVVLPKSNVPDRIKTRKIFTLST 282
 QY 287 ADFAINAVAK 297
 Db 283 EDFAINNISK 293

RESULT 4
 AAW29217
 ID AAW29217 standard; protein; 312 AA.
 AC AAW29217;
 XX
 DT 02-MAR-1998 (first entry)
 XX
 DE S. cerevisiae uronate dehydrogenase.
 XX
 KW uronate reductase; uronate dehydrogenase; Saccharomyces cerevisiae;
 KW synthesis; L-ascorbic acid.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN DE19604798-A1.
 XX
 PD 14-AUG-1997.
 XX
 PF 09-FEB-1996; 96DE-01004798.
 XX
 PR 09-FEB-1996; 96DE-01004798.
 XX
 PA (HERB-) HERBSTREITH & FOX PEKTIN-FAB NEUENBUERG.

PI Mattes R, Kulbe K;
 XX
 DR WPI; 1997-403701/38.
 DR N-PSDB; AAT87004.
 XX
 PT Recombinant uronate reductase and L-gulono-gamma-lactone oxidase - used
 PT in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 XX
 XX Claim 27; Page 18-19; 26pp; German.
 XX
 CC The present sequence represents a uronate dehydrogenase (reductase)
 CC isolated from *Saccharomyces cerevisiae*, which is used for synthesis of L-
 CC ascorbic acid. DNA encoding L-gulono-gamma-lactone oxidase (GluOx) is
 CC also used in the production of L-ascorbic acid. The uronate dehydrogenase
 CC has higher affinity for substrate (D-glucuronic or galacturonic acids)
 CC and coenzymes than the similar enzyme from *Lipomyces starkeyi*, also
 CC better stability and greater yield when produced recombinantly
 XX
 XX Sequence 312 AA;
 SQ

Query Match 33.2%; Score 574.5; DB 2; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.9e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

QY 5 KTFTLSNGVKIPGVGFTFASGSKGETYTAVTALKTGYRHLDCAWYVINEGEVGEIR 64
 Db 10 KILSLNTGAQIPQIGLGTWQSK--ENDAYKAVLTALKDGYRHIDTAIYRNEDQVGOAIK 67
 QY 65 DFLKENSVPKREDIFVCTKWNHLHRYEDVLMISDLSKRLGLDYDMFLVHWPAAEKN 124
 Db 68 D-----SGVPREEIFVTIKLWCTQHHEPEV--ALDQSKRLGLDYDVLVHWPAAEL--- 117
 QY 125 GQGEPKIGPDKGVILKDLTENPE-----PTWRAMEKIVEDRKARSIGV 168
 Db 118 -----DPAYIKNEDILSVPTKDGSRADITNWNFIKTWELMQELPKTKTKAVGV 168
 QY 169 SNWTIADLEKM--SKFAKVMPHANQIEIHPPLNEELVOYCFCKNIMPVAYSPLGSONOV 226
 Db 169 SNFSINNKLKOLLASQGNKLTAAANQVEIHPLLPQDELINFCCKSGIWEAYSPLGS---- 224
 QY 227 PTTGERSVENKTLNIEAEKGNLTAAQVLIAGLRGYYVLPKSNPKRIESNFKSIELSD 286
 Db 225 --TDAPLKEPVEILEIAKNNVQPGHVVISHWVQGVVLPKSNVPDRIKTRKIFTLST 282
 QY 287 ADFAINAVAK 297
 Db 283 EDFAINNISK 293

RESULT 5
 AAG63565
 ID AAG63565 standard; protein; 312 AA.
 XX
 AC AAG63565;
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE Amino acid sequence of a nuclear gene encoding a yeast protein.
 XX
 KW Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO200155342-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US003186.
 XX
 PR 31-JAN-2000; 2000US-00494921.
 PR 08-DEC-2000; 2000US-00734237.
 XX

```

Query Match      100.0%; Score 1731; DB 5; Length 325;
Best Local Similarity 100.0%; Pred. No. 5.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNGKTFTLSNGVKIPGVGFTFASGSKGETYTAVTALTGTGVRHLDCAWYLNVEGVG 60
DB 1 MSNGKTFTLSNGVKIPGVGFTFASGSKGETYTAVTALTGTGVRHLDCAWYLNVEGVG 60
QY 61 EGIRDFLKENPSVKREDIFVCTKWNHLHRYEDVLWSIDDSLKRLGLDYDMFLVHWPIA 120
DB 61 EGIRDFLKENPSVKREDIFVCTKWNHLHRYEDVLWSIDDSLKRLGLDYDMFLVHWPIA 120
QY 121 AEKNGQGEPKIPGDKYVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
DB 121 AEKNGQGEPKIPGDKYVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
QY 181 KFAKVMPHANQIEIHPFLPNEELVQYCFSKNIMPVAYSPLGSONQVPTTGERVSENKTLN 240
DB 181 KFAKVMPHANQIEIHPFLPNEELVQYCFSKNIMPVAYSPLGSONQVPTTGERVSENKTLN 240
QY 241 EIAEKGNTLAQVLIANGRLRGYVVLPKSSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
DB 241 EIAEKGNTLAQVLIANGRLRGYVVLPKSSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
QY 301 FRFVNMDTFGYDVPBETAKNLSA 325
DB 301 FRFVNMDTFGYDVPBETAKNLSA 325

RESULT 2
ADE39629
ID ADE39629 standard; protein; 325 AA.
AC ADE39629;
XX
DT 29-JAN-2004 (first entry)
XX
DE Penicillium reducing enzyme protein, SEQ ID No 3.
XX
KW optically active 2-hydroxycycloalkanecarboxylic acid ester;
KW 2-oxocycloalkanecarboxylic acid ester; enzyme; penicillium; reduction.
XX
OS Penicillium citrinum.
XX
FN EPI323827-A2.
XX
PD 02-JUL-2003.
XX
PF 20-DEC-2002; 2002EP-00258814.
XX
PR 27-DEC-2001; 2001JP-00395884.
PR 27-DEC-2001; 2001JP-00395885.
PR 10-APR-2002; 2002JP-00107648.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Asako H, Wakita R, Itoh N;
XX
DR WPI; 2003-723302/69.
DR N-PSDB; ADE39630.
XX
PT Producing optically active 2-hydroxycycloalkanecarboxylic acid ester by
PT reacting 2-oxocycloalkanecarboxylic acid ester with transformant having
PT ability to asymmetrically reduce the ester to optically active ester.
XX
PS Claim 9; SEQ ID NO 3; 45pp; English.
XX
CC The invention relates to a novel method for producing an optically active
CC 2-hydroxycycloalkanecarboxylic acid ester. The novel method involves:
CC allowing 2-oxocycloalkanecarboxylic acid ester to react with a
CC transformant, or a dead cell or their extract, artificially provided with
CC the ability to asymmetrically reduce a 2-oxocycloalkanecarboxylic acid

```

```

ester to the optically active 2-hydroxycycloalkanecarboxylic acid; and an
ability to regenerate a coenzyme on which an enzyme having the above
ability depends; and collecting the resulting optically active 2-
hydroxycycloalkanecarboxylic acid. The optically active 2-
hydroxycycloalkanecarboxylic acid is useful as an intermediate for the
production of bioactive substances. This sequence represents a
penicillium enzyme protein used to reduce 2-oxocycloalkanecarboxylic acid
ester to the optically active 2-hydroxycycloalkanecarboxylic acid as part
of the method of the invention.
XX
SQ Sequence 325 AA;

Query Match      100.0%; Score 1731; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 5.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNGKTFTLSNGVKIPGVGFTFASGSKGETYTAVTALTGTGVRHLDCAWYLNVEGVG 60
DB 1 MSNGKTFTLSNGVKIPGVGFTFASGSKGETYTAVTALTGTGVRHLDCAWYLNVEGVG 60
QY 61 EGIRDFLKENPSVKREDIFVCTKWNHLHRYEDVLWSIDDSLKRLGLDYDMFLVHWPIA 120
DB 61 EGIRDFLKENPSVKREDIFVCTKWNHLHRYEDVLWSIDDSLKRLGLDYDMFLVHWPIA 120
QY 121 AEKNGQGEPKIPGDKYVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
DB 121 AEKNGQGEPKIPGDKYVILKDLTENPEPTWRAMEKIYEDRKARSIGVSNWTIADLEKMS 180
QY 181 KFAKVMPHANQIEIHPFLPNEELVQYCFSKNIMPVAYSPLGSONQVPTTGERVSENKTLN 240
DB 181 KFAKVMPHANQIEIHPFLPNEELVQYCFSKNIMPVAYSPLGSONQVPTTGERVSENKTLN 240
QY 241 EIAEKGNTLAQVLIANGRLRGYVVLPKSSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
DB 241 EIAEKGNTLAQVLIANGRLRGYVVLPKSSNPKRIESNFKSIELSDADFEAINAVAKGRH 300
QY 301 FRFVNMDTFGYDVPBETAKNLSA 325
DB 301 FRFVNMDTFGYDVPBETAKNLSA 325

RESULT 3
AAW29220
ID AAW29220 standard; protein; 312 AA.
XX
AC AAW29220;
XX
DT 02-MAR-1998 (first entry)
XX
DE S. cerevisiae uronate dehydrogenase variant Q29G.
XX
KW uronate reductase; uronate dehydrogenase; Saccharomyces cerevisiae;
KW synthesis; L-ascorbic acid; variant.
XX
OS Saccharomyces cerevisiae.
XX
FH Key Location/Qualifiers
FT Misc-difference 29
FT /label= Q29G
FT /note= "wild type glutamine replaced by glycine"
XX
FN DE19604798-A1.
XX
PD 14-AUG-1997.
XX
PF 09-FEB-1996; 96DE-01004798.
XX
PR 09-FEB-1996; 96DE-01004798.
XX
PA (HERB-) HERBSTREITH & FOX PEKTI-FAB NEUENBUERG.
PI Mattes R, Kulbe K;
XX

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 13:58:21 ; Search time 60 Seconds
(without alignments)
1530.466 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1731

Sequence: 1 MNGKFTLNSGVKIPGVCF.....MKDTFGYDWPETAKNLSA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1731	100.0	325	5	Abp77965 Protein w
2	1731	100.0	325	7	Ade39629 Penicilli
3	575.5	33.2	312	2	Aaw29220 S. cerevi
4	574.5	33.2	312	2	Aaw29217 S. cerevi
5	574.5	33.2	312	4	Aag63565 Amino aci
6	574.5	33.2	312	5	Abg93198 S. cerevi
7	574.5	33.2	313	4	Aag63566 Synthetic
8	573.5	33.1	312	2	Aaw29218 S. cerevi
9	565.5	32.7	312	2	Aaw29219 S. cerevi
10	559.5	32.3	323	2	Aar96294 Carboxyl
11	555	32.1	313	4	Aag63564 Amino aci
12	550	31.8	295	5	Abg93303 C. albica
13	547.5	31.6	344	5	Abg93050 S. cerevi
14	547.5	31.6	344	5	Aau76345 Yeast D-a
15	547	31.6	328	5	Abp53552 T. megach
16	545.5	31.5	328	5	Abp53551 T. megach
17	543.5	31.4	313	4	Aab47466 G. max al
18	543.5	31.4	313	7	Aae39522 Soybean a
19	541	31.3	322	4	Aag63561 Amino aci
20	538	31.1	330	5	Abp53550 T. megach
21	533.5	30.8	309	3	Aag07671 Arabidops
22	529.5	30.6	309	3	Aag33086 Arabidops
23	529.5	30.6	309	3	Aag50693 Arabidops
24	529	30.6	290	3	Aag31845 Arabidops
25	529	30.6	315	3	Aag40418 Arabidops

26	528	30.5	280	5	Abp30416 Streptoco
27	528	30.5	301	5	Abp27439 Streptoco
28	527.5	30.5	302	4	Abb62062 Drosophil
29	527.5	30.5	309	3	Aag43890 Arabidops
30	525	30.3	279	6	Abu18432 Protein e
31	523.5	30.2	313	2	Aay06231 Alfalfa a
32	523.5	30.2	311	3	Aag17752 Arabidops
33	522.5	30.2	311	3	Aag40415 Arabidops
34	522.5	30.2	311	3	Aag04944 Arabidops
35	521	30.1	300	6	Abu24848 Protein e
36	519.5	30.0	311	3	Aag44797 Zea mays
37	519.5	30.0	364	3	Aab47468 Z. mays N
38	518.5	30.0	308	4	Aae39524 Corn NADP
39	518.5	30.0	308	7	Aag09538 Arabidops
40	518.5	30.0	309	3	Abb65274 Drosophil
41	516.5	29.8	320	4	Abb65274 Drosophil
42	516	29.8	279	6	Abu17544 Protein e
43	516	29.8	280	6	Abu01927 S. pneumo
44	516	29.8	280	6	Abu46138 Protein e
45	516	29.8	281	3	Aay81549 Streptoco

ALIGNMENTS

RESULT 1

ABP77965

ID ABB77965 standard; protein; 325 AA.

XX

AC ABB77965;

XX

DT 22-OCT-2002 (first entry)

XX

DE Protein which is capable of producing (S)-4-bromo-3-hydroxybutanoate.

XX

KW (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutic;
agrochemical; 4-cyano-3-hydroxybutanoic acid.

XX

OS Penicillium citrinum.

XX

PN EP1213354-A2.

XX

PD 12-JUN-2002.

XX

PF 07-DEC-2001; 2001EP-00310251.

XX

PR 07-DEC-2000; 2000JP-00372704.

PR

PR 15-JAN-2001; 2001JP-00006144.

PR

PR 02-FEB-2001; 2001JP-00026594.

PR

PR 11-JUN-2001; 2001JP-00175175.

XX

(SUMO) SUMITOMO CHEM CO LTD.

PI

Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;

XX

WPI: 2002-550350/59.

DR

N-PSDB; ABL59376, ABL59397, ABL59398.

DR

New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by

asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals

and agrochemicals.

PT

Claim 14; Page 29-31; 56pp; English.

XX

The present sequence represents a Penicillium citrinum protein which is

capable of producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically

reducing 4-bromo-3-oxobutanoate. The protein and polynucleotides are

useful for producing optically active (S)-4-bromo-3-hydroxybutanoate,

which is useful as an intermediate in the production of pharmaceuticals

and agrochemicals. The (S)-4-bromo-3-hydroxybutanoate produced can also

be used to produce 4-cyano-3-hydroxybutanoic acid

XX

SQ Sequence 325 AA;